Supplementary materials to Optimal proteome allocation and the evolution of cross-feeding

Florian Labourel[1], Frédéric Menu[1], Vincent Daubin[1], and Etienne Rajon[1]

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

December 16, 2021

Text S1 Total optimal content and cellular constraints

Numerous constraints may affect the total optimal content of cells. It was shown by Dill et al. [2011] that the amount of proteins should establish at an optimal intermediate level due to the deleterious impact an extra expression of proteins would have on diffusion. However, they did not account for the protein burden of producing these molecules. Their focus was besides on the total content without regard to the specific allocation of this content. This matters because a cell should be more prone to invest in a task when it entails large increases of fitness, and/or when these tasks are necessary to survive. Here, we assessed the effects of cellular constraints asumed to be within documented ranges [Wagner, 2005, Chou et al., 2014, Lynch and Marinov, 2015, Kafri et al., 2016, Blanco et al., 2018, Andrews, 2020].

We modelled a pathway initiated by a glucose carrier protein that facilitates diffusion, whose features correspond to average values - $V_{Tm} = 1mM/s$, $K_T = 10mM, \alpha = 1$ - for those reported in yeasts [Teusink et al., 1998, Maier et al., 2002]; notice that we do not report results of the influence of transporters since it had none on the processes we are interested in. The chemical equation for facilitated transport can be approached by the following equation [ter Kuile and Cook, 1994, Bosdriesz et al., 2018]:

$$\frac{d[S_{\text{in}}]}{dt} = V_{Tm}.\frac{[S_{\text{out}}] - [S_{\text{in}}]}{K_T + ([S_{\text{out}}] + [S_{\text{in}}]) + \alpha.\frac{[S_{\text{out}}][S_{\text{in}}]}{K_T}} \tag{S1}$$

To match with central carbon metabolism, composed by glycolysis and tricarboxylic acid cycle, the pathway modelled is comprised of 40 enzymes that obey Michaelis Menten kinetics, according to the following scheme (we relax the absence of reversibility later, by using Briggs-Haldane equations [Briggs and Haldane, 1925]):

$$\mathrm{E}_i + \mathrm{S}_i \underset{\mathrm{k_r}}{\overset{\mathrm{k_{f,act}}}{\rightleftharpoons}} \mathrm{ES}_i \xrightarrow{\mathrm{k_{cat}}} \mathrm{E}_i + \mathrm{P}_i, \tag{S2}$$

where $k_{f,act}$ is the *in vivo* value of k_f when accounting for the influence of crowding on difusive processes. This influence is modelled by the following equation, already justified elsewhere [Labourel and Rajon, 2021]:

$$k_{f,act} = k_f.10^{-([E_{basal}]+\sum_{i=1}^{40}[E_{tot,i}])/[M_b]},$$

where $[E_{tot,i}] = [E_i] + [ES_i]$,$[M_b] = 3 \cdot 10^{-3}M$ represents the scaling factor for the effect of diffusion, while E_{basal} denotes the constant amount of protein allocated to other tasks than those of central carbon metabolism. Notice that compared with the previous reference, where we were mostly interested in setting a qualitatively realistic crowding limit, we have here refined this estimate to get as close as possible from physical findings [Blanco et al., 2018, Andrews, 2020] and from a realistic cellular protein fraction [Ellis, 2001, Dill et al., 2011]. Noticeably, the effect appears somewhat higher than previously, reflecting the need of an increase in the expression of cellular machineries such as ribosomes [Klumpp et al., 2013, Kafri et al., 2016] in order to produce more proteins.

k_f, k_{cat} and k_r are first set to values in line with high estimates for central carbon metabolism [Bar-Even et al., 2011].

The selective pressure can be approximated by a linear degradation parameter η_d competing with enzymes for their substrate (often denoted as a product since it also coincides with the product of the previous reaction), according to the following scheme:

$$\begin{array}{ccccc} \mathrm{P}_i + \mathrm{E}_{i+1} & \underset{k_r}{\overset{k_f}{\rightleftharpoons}} & \mathrm{E}_{i+1}\mathrm{P}_i & \xrightarrow{k_{\mathrm{cat}}} & \mathrm{E}_{i+1} + \mathrm{P}_{i+1}, \\ \downarrow \eta_d & & & & \\ \mathrm{P}_{i,out} & & & & \end{array} \tag{S3}$$

where P_i is the product of the previous enzymatic reaction and E_{i+1} is the focal enzyme of the pathway, specialised at processing this product. As previously stated, the degradation rate η_d that competes for metabolites applies to each reaction in the pathway.

In the first section below, we tested the influence of the selective pressure imposed by the degradation rate under different assumptions. As mentioned in the manuscript, our Adaptive Dynamics model relies on the dynamic competition between cells, and, more specifically, on the capability of mutants to invade resident strategies when these mutants are rare.

Text S1.1 Influence of the degradation rate on optimal enzyme concentration

Because Adaptive Dynamics rely on the resident strategy having reached its ecological equilibrium, it is necessary to set an amount of net energy produced - the energy produced minus expenses entailed by protein production - that corresponds to this equilibrium in which births compensate for deaths. We set this amount to $\Phi_{eq} = 10^{-4} M.s^{-1}$. This parameter only influences the relative cost dedicated to the sustainment of the proteome. Besides, there is also a need to set the size of cells - $r_c = 1\mu m$ corresponding to a volume of $V_c \approx 4.2\mu m^3$ (roughly that of *E.coli*) - and to consider a specific fraction of the environment - set to $V_{env} = 1000\mu m^3$ without cells (it is the volume of the environment that is free of cells). Notice that the size of cells matter when we study the influence of permeability for passive diffusion depends on the SA:V ratio, which decreases when cells are bigger - this is discussed in the section about cross-feeding. On the contary, the size of the environment does not matter: it only modifies the number of cells coinciding with the ecological equilibrium, which, in Adaptive Dynamics, is not involved in the outcome due to genetic drift not being considered. Finally, the "chemostats" parameters are set such that the enrichment rate equals the dilution rate, $\alpha = 10^{-3}$ and $\beta = 10^{-3}$ yielding a steady-state concentration in a cell depleted medium of $[S_{out}]^* = 1M$. These coefficients are in line with estimates for diffusion coefficients of metabolites in solvant. Lowering them changes the speed at which nutrients are brought to the environment, and, therefore, influences how possible it is for cells to thrive (because they need to sustain a given flux, that is a given amount of production per unit time, which is partly dependent on how quickly the environment is replenished). The set of parameters that are fixed at this stage is summed up in the following table:

Parameters	$\Phi_{eq}(M^{-1}s^{-1})$	$V_c(\mu m^3)$	$V_{env}(\mu m^3)$	$\alpha(Ms^{-1})$	$\beta(s^{-1})$
Values	10^{-4}	4.2	1000	10^{-3}	10^{-3}

Table S1: Set of constant parameters used to simulate competition

As the system yields an ecological equilibrium that needs be solved numerically through a two step process - see Materials & Methods of the article for details - it is not possible to determine the joint influence of the whole set of parameters. Instead, we varied them on a pairwise basis where the degradation rate is always the focal variable while other parameters are all set but one (or two, if necessary). When not explicitly mentioned, these parameters are set according to the following table:

Parameters	$k_f(M^{-1}s^{-1})$	$k_{cat}(s^{-1})$	c	E_{basal}
Values	10^7	$10^{2.5}$	$10^{-2.5}$	$5.5 \cdot 10^{-3}$

Table S2: Set of basic settings used to determine their individual influence

Notice also that in this first subsection, k_r is always set to equal k_{cat}.

Text S1.1.1 Influence of enzyme kinetic parameters and the concentration of the first enzyme of the pathway

To evaluate how kinetic parameters impact the optimal cell content, we modelled the process for three different values of enzyme efficiency, varying them by half an order of magnitude. This is made necessary because kinetic parameters define another parameter on which Natural Selection can act to promote enzyme activity. But there is an essential difference between those two: while enzyme levels should be easily tunable, since they are constrained by their consequence on the working of cells, kinetic parameters, for their part, are constrained by the inherent difficulty to improve due to the scarcity and findability, if existing, of very efficient phenotypes. Simulatenously, there is also another factor impacting the fitness landscape on which cells evolve and that in turn modifies the optimal content, which is the concentration of the most upstream enzyme. Indeed, this enzyme determines how much nutrients is taken from the environment at each timestep, and therefore, the flux of substrate entering the metabolic pathway. Because facilitated diffusion relies on carrier proteins and the substrate gradient along the cell membrane, it has its own selective pressure, potentially differing from downstream enzymes [ter Kuile and Cook, 1994, Labourel and Rajon, 2021].

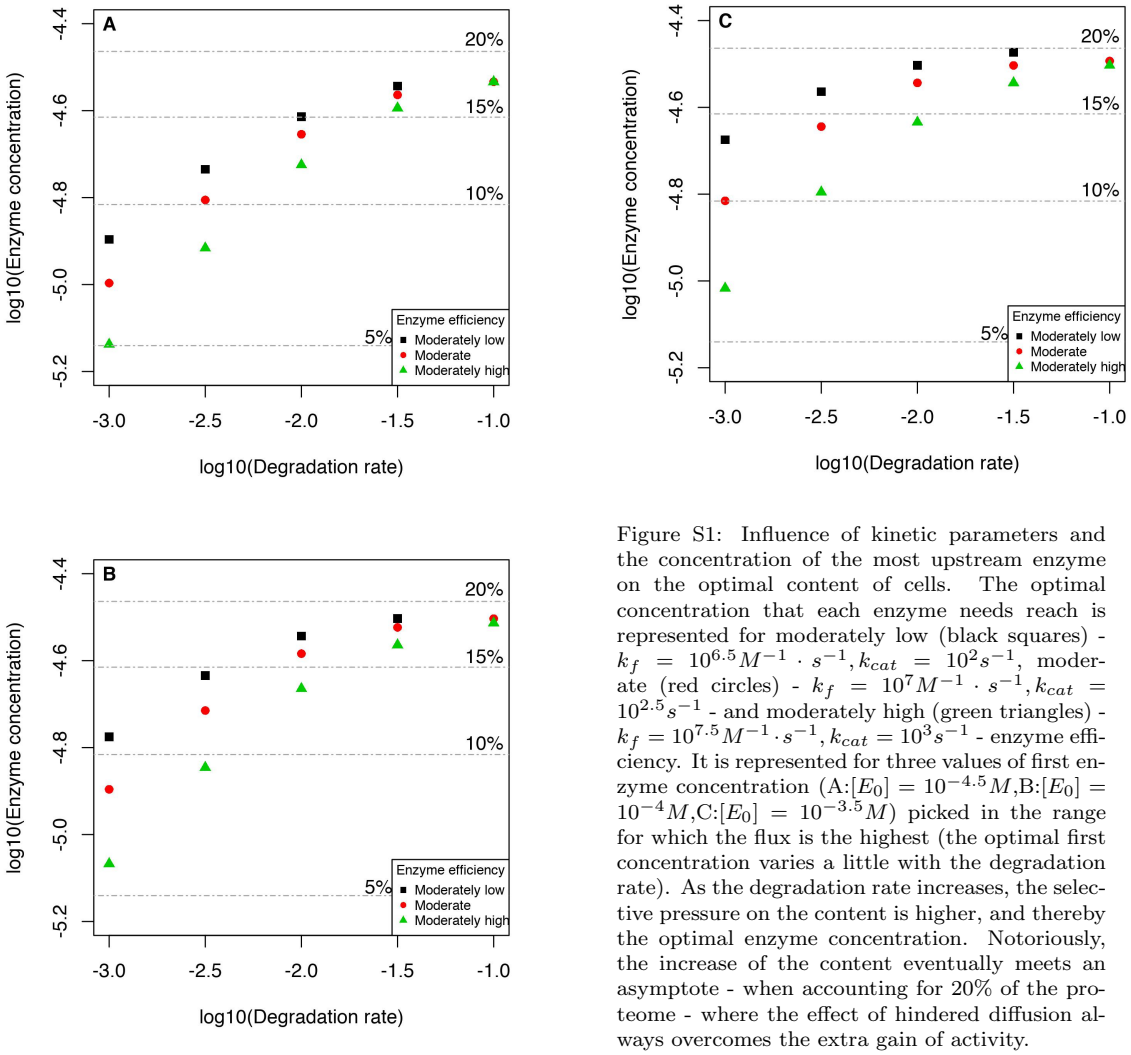

Figure S1: Influence of kinetic parameters and the concentration of the most upstream enzyme on the optimal content of cells. The optimal concentration that each enzyme needs reach is represented for moderately low (black squares) - $k_f = 10^{6.5} M^{-1} \cdot s^{-1}, k_{cat} = 10^2 s^{-1}$, moderate (red circles) - $k_f = 10^7 M^{-1} \cdot s^{-1}, k_{cat} = 10^{2.5} s^{-1}$ - and moderately high (green triangles) - $k_f = 10^{7.5} M^{-1} \cdot s^{-1}, k_{cat} = 10^3 s^{-1}$ - enzyme efficiency. It is represented for three values of first enzyme concentration (A:$[E_0] = 10^{-4.5} M$,B:$[E_0] = 10^{-4} M$,C:$[E_0] = 10^{-3.5} M$) picked in the range for which the flux is the highest (the optimal first concentration varies a little with the degradation rate). As the degradation rate increases, the selective pressure on the content is higher, and thereby the optimal enzyme concentration. Notoriously, the increase of the content eventually meets an asymptote - when accounting for 20% of the proteome - where the effect of hindered diffusion always overcomes the extra gain of activity.

The higher the concentration, the higher the nutrient gradient, and the higher the flux, as long as the expression of downstream enzymes do not impede too much cellular diffusion of macromolecules. This means that a more highly expressed upstream enzyme increases the selective pressure on downstream enzymes, which increases their optimal expression level and eventually up to a point where the effect of an increment of expression is deleterious. Inasmuch as the expression of the first enzyme is part of the hindering process

and in spite of the previous argument, its occurrence may also decrease the optimal concentration for other enzymes. This latter phenomenon can be observed marginally on Figure S1 - C, where the optimal concentration decreases more for a higher degradation rate - compare C with B, for the highest degradation rates. Notice that the A situation corresponds to the example pf PIPs shown in Figure S14 of Appendix. For high values of degradation rates (around - that are consistent with the selective pressure observed for enzymes involved in central carbon metabolism [Bar-Even et al., 2011, Labourel and Rajon, 2021], the optimal content approximately converges towards an asymptotic value of 20%. Above this level, the extra activity does not offset the loss of intracellular diffusibility, a conclusion that holds qualitatively for all enzyme efficiencies and first enzyme concentration. Notice that this degradation rate is on another hand inconsistent owing to the amount of nutrients lost in the process, a limit on which we later elaborate.

Text S1.1.2 Influence of other cellular parameters

As to study the influence of other parameters, we then set $k_f = 10^7 M^{-1} s^{-1}$, $k_{cat} = 10^{2.5} s^{-1}$ and $k_r = k_{cat}$, approximately an order of magnitude higher than median estimates of Bar-Even et al. [2011]. In parallel, the concentraiton of the upstream enzyme is set to $0.1mM$. In the following, we tested the influence of other parameters and demonstrated that the main influence is set, as expected, by the degradation rate.

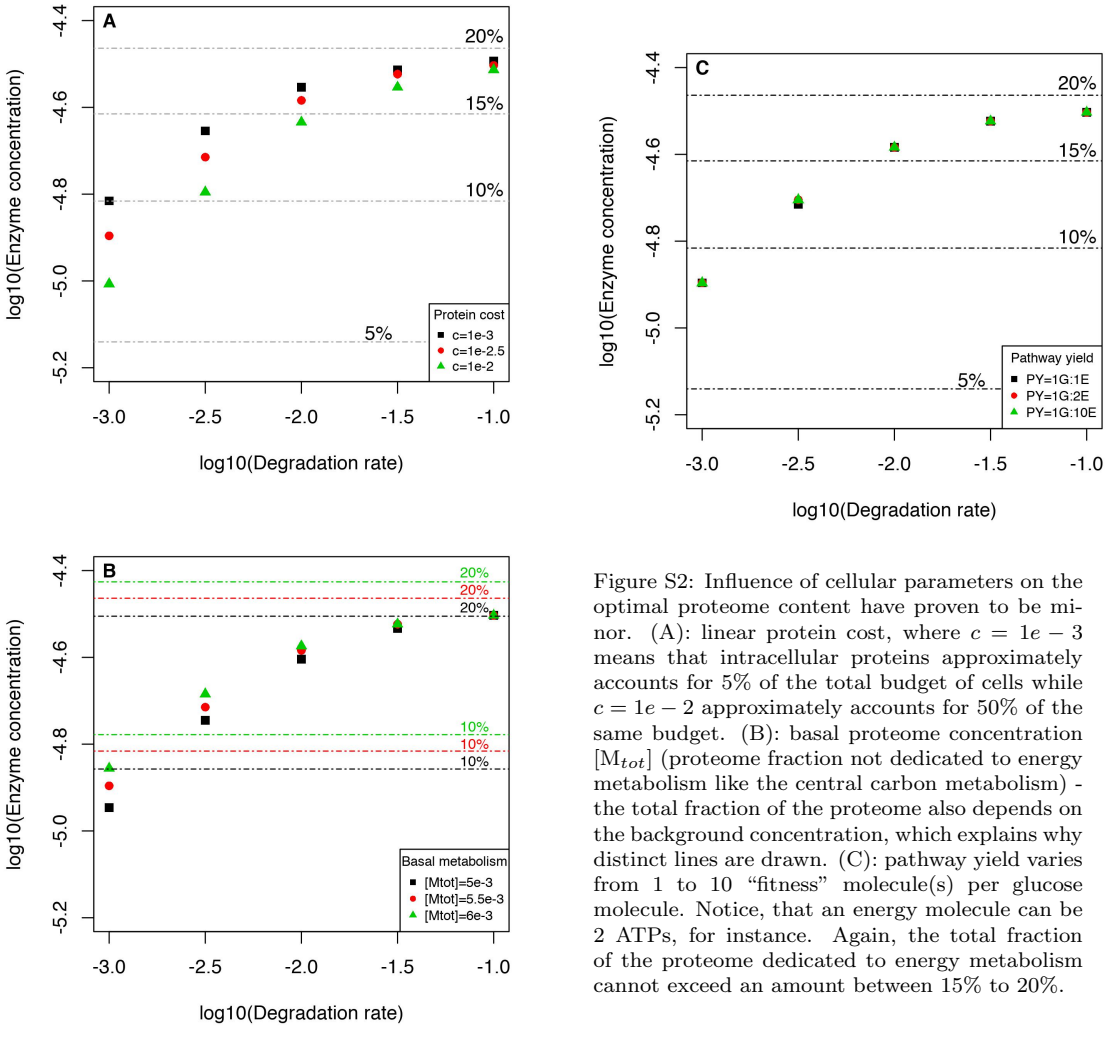

Figure S2: Influence of cellular parameters on the optimal proteome content have proven to be minor. (A): linear protein cost, where $c = 1e-3$ means that intracellular proteins approximately accounts for 5% of the total budget of cells while $c = 1e-2$ approximately accounts for 50% of the same budget. (B): basal proteome concentration [M_{tot}] (proteome fraction not dedicated to energy metabolism like the central carbon metabolism) - the total fraction of the proteome also depends on the background concentration, which explains why distinct lines are drawn. (C): pathway yield varies from 1 to 10 "fitness" molecule(s) per glucose molecule. Notice, that an energy molecule can be 2 ATPs, for instance. Again, the total fraction of the proteome dedicated to energy metabolism cannot exceed an amount between 15% to 20%.

Text S1.1.3 Influence of environment parameters

The replenishment of the environment results from the flux parameter α while the degradation results from the rate parameter β. They had no impact on the optimal content, only changing the demographic equilibrium (with a lower α, the steady-state population diminishes and may even vanish, even if the steady-state

concentration in the environment is high). We do not report these results but the scripts that generated the results are available in the repository.

Text S1.2 Influence of the degradation rate on the concentrations of metabolites along the pathway

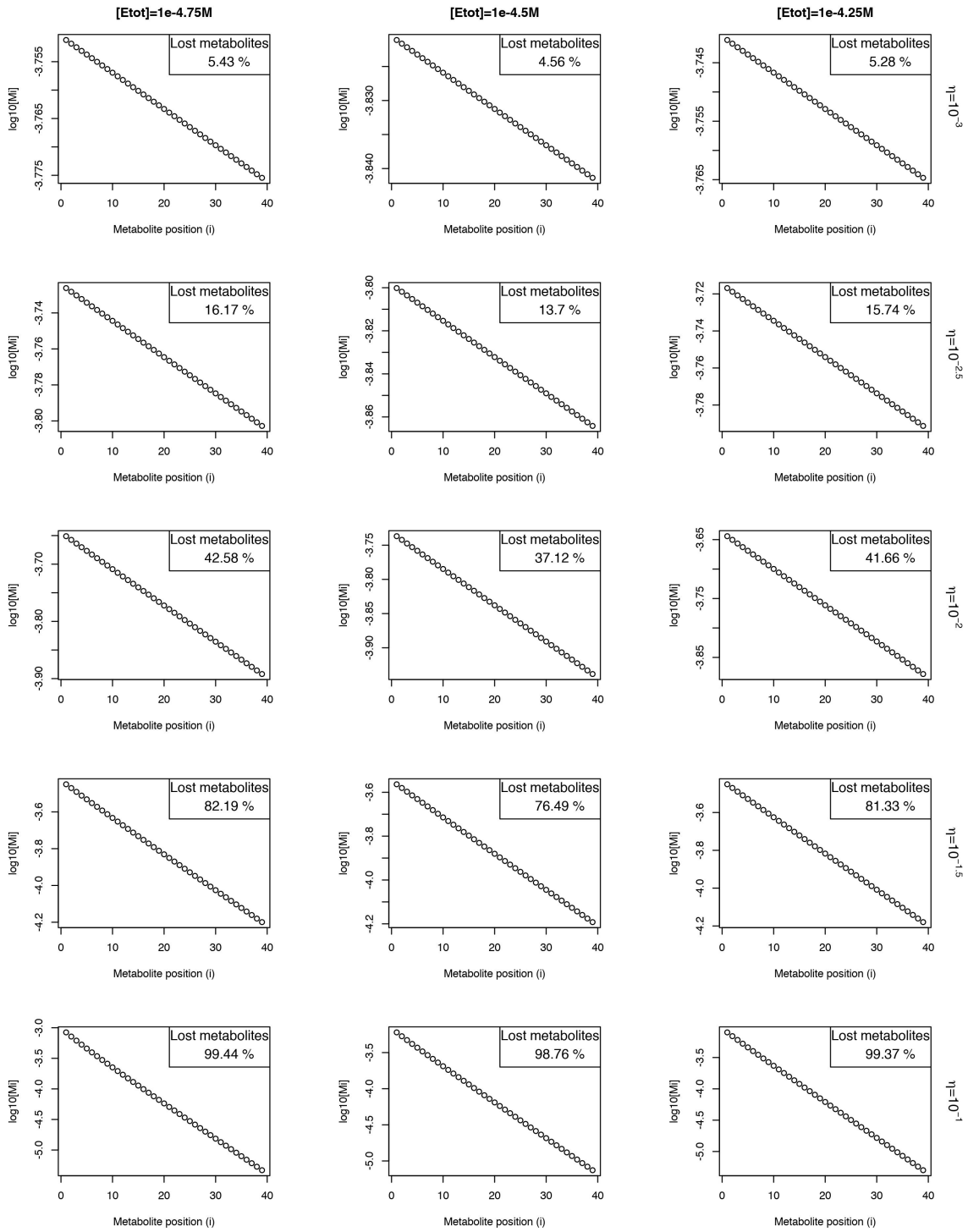

Figure S3: Metabolite losses along the pathway due to various degradation rates (rows) considering different enzyme concentrations (columns): the loss is log-linear in any case. The first two lines ($\eta < 10^{-2}s^{-1}$) coincide with realistic losses of metabolites. However, the selective pressure acting on enzymes was shown to correspond to the highest degradation rates ($\eta > 10^{-2}$), raising a limit to approach the process through the sole degradation rate.

In this section, we determined the effect of the degradation rate on the loss of metabolites along the pathway. Indeed, each reaction copes with this linear effect that accumulate progressively. Assuming degradation rates in line with estimates accounting for reversibility yields high losses, and a high decrease of flux, which is unrealistic. This may bias the estimation of the optimal concentration insofar as the concentration of enzyme should not be equally spread along the pathway, contrary to what we model. This even truer for the case where we study the allocation between two parts of a pathway, one being downstream and following the last reaction of the one, which is the upstream part. To overcome these limits and test how they may bias outcomes, we also studied a model including metabolite toxicity and reversibility within the pathway.

Text S1.3 Influence of more realistic set of constraints on the optimal content

We determined the effect of toxicity using a non-linear effect of toxicity - as in Chou et al. [2014] - affecting fitness according to the following equation:

$$f = (\Phi - cost \cdot \sum_{i=1}^{40}[E_{tot,i}]) \times \frac{T}{T + \sum_{i=1}^{40}[M_i]},$$

where T is a toxicity constant, which, when reached by the total metabolite content, cut fitness by half.

For example, a toxicity constant set to $10^{-1}M$ means that if the sum of all 40 metabolites involved in the pathway equals this amount, then fitness is half what it would have been without this constraint. Toxicity has a similar qualitative effect than those studied previously - see Figure S9, where T varies from $10^{-}2M$ to $1M$. Nevertheless, its impact on the optimal content is larger, as a high toxicity alone is sufficient to induce the same selective pressure than a high degradation rate. The maximal optimal content – coinciding with approximately 20% of the pathway dedicated to the energy metabolism – does not differ from the ones yielded by other parameters.

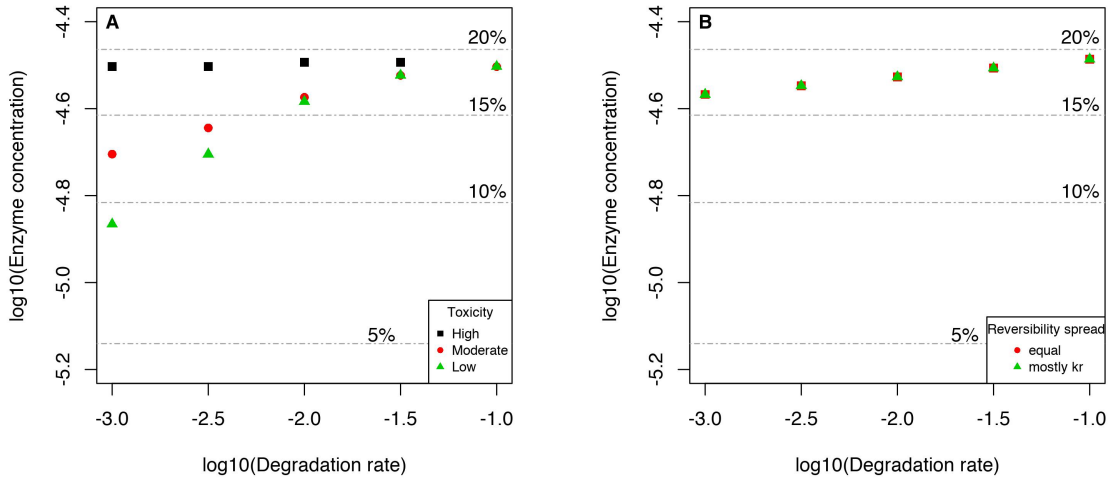

Figure S4: Influence of metabolite toxicity (A) and reversibility (B, combined with a basal toxicity) on the optimal content. Toxicity T – High: $T = 10^{-2}M$; Moderate: $T = 10^{-1}M$; Low: $T = 1M$ – drastically increases the selective pressure acting on enzyme concentration (see the high toxicity - black points - in A), so that even with a low degradation rate, the optimal cell content is largely increased. With a high toxic level of metabolites, the selective pressure on enzyme concentration is no longer influenced by the degradation process except that the higher it is, the lower the efficiency for any strategy (in terms of expression). Assuming a moderate toxicity – $K_{rev} = 1/9$ – in (B) that combines with a realistic level for reaction reversibility, the selective pressure is again less dependent on the degradation rate, although changes in where the reversiblity is located – "equal": $k_r = k_{cat}/3$, $k_{inh} = k_f/3$; "mostly k_r" :$k_r = k_{cat}$, $k_{inh} = k_f/9$ – only impacts the viability of cells (when relying mostly on k_{inh}, cells are no longer viable - results not displayed here).

Reversible reactions obey the following scheme (where (i) denotes the i^{eth} reaction):

$$\mathrm{S}_i + \mathrm{E}_i \underset{k_{r,i}}{\overset{k_{f,i}}{\rightleftharpoons}} \mathrm{E}_i S \underset{k_{inh,i}}{\overset{k_{\mathrm{cat},i}}{\rightleftharpoons}} \mathrm{E}_i + \mathrm{P}_i \tag{S4}$$

Reversibility was therefore studied by considering that it can affect either k_r, k_{inh} or both. The level of reversibility was set to a specific value ($K_{rev} = 1/K_{eq} = [S]_{eq}/[P]_{eq} = 1/9$), which is the geometric mean of that for reacions involved in the central carbon metabolism and whose value have been summarised in Li et al. [2011]. As reversibility can be spread between two parameters, it is necessary to see how the flux reacts to this intrinsic process under the cellular constraints. Notice that we simplify this process by assuming it does not evolve, even though it was shown that organisms should in principle optimise the energy profile determining how reversibility is spread [Heinrich et al., 1991, Klipp and Heinrich, 1994]. If the reversibility impedes only the parameter k_{inh}, there is no degradation rate susceptible to produce a flux high enough to compensate for the need to sustain its pool of proteins. Therefore, we report results only for the two cases where reversibility acts only on k_{inh}, or that for which it is equally spread between both parameters ($k_{cat} = k_r/3$ and $k_{inh} = k_f/3$). The effect of reversibility also increases largely the selective pressure, although for low degradation rates, there is stil a little room for extra protein expression. Notice that we do not report the influence of reversibility solely, for it proved to be similar.

Text S2 Differential allocation between subparts of pathways: a toy model

In this section, we introduce a toy model designed to unraval the intricacies behind the optimal allocation strategy along a pathway. Instead of considering a long pathway, we focus on a pathway made up of two consecutive reactions that contribute to fitness, where the flux prior to the first reaction is denoted by Φ_0. Based on insights from the flux control theory [Kacser and Burns, 1973, Heinrich and Rapoport, 1974] and a more recent mechanistic approach [Labourel and Rajon, 2021] partly relaxing the need for unsaturated reactions, the flux sustained by enzymes of reaction (i) can be written as:

$$\Phi_i = \Phi_{i-1} \frac{[E_i]}{K + [E_i]}, \tag{S5}$$

where $[E_i]$ denotes the total concentration of enzyme (i) and K represents a phenomenological saturation parameter acting on enzymes and involving different constraints emerging within a pathway [Hartl et al., 1985, Kaltenbach and Tokuriki, 2014].

With two reactions in pathway, the system can be summarised as follows:

$$\begin{cases} \Phi_1 = \Phi_0 \dfrac{[E_1]}{K + [E_1]} & \text{(S6a)} \\ \Phi_2 = \Phi_1 \dfrac{[E_2]}{K + [E_2]} & \text{(S6b)} \\ W = \Phi_1 + \Phi_2 - c \cdot ([E_1] + [E_2]) & \text{(S6c)} \end{cases}$$

with c representing the cost of protein production and Φ_1, Phi_2 the fluxes that directly (and equally) contribute to fitness.

Making the simplistic assumption that the parameter K is identical for both reactions, fitness can therefore be written as:

$$W = \Phi_0 \cdot \left(\frac{[E_1]}{[E_1] + K}\right)\left(1 + \frac{[E_2]}{[E_2] + K}\right) - c \cdot ([E_1] + [E_2])$$

Note that what generates the flux Φ_0 does not matter for our purpose, though it may be seen as the flux produced by carrier proteins transporting a specific nutrient.

The optimal allocation stemming from such a system is reached when the extra fitness gained by increasing either one of the concentration equals that obtained with the other one. Indeed, at the point where any increase of the total concentration does not entail any extra fitness, this concentration has to be spread between pathways and if it increases fitness to increase the allocation in one pathway, it implies that there is a corollay interest to decrease the allocation to the other one. This condition can be written as:

Condition 1 $\frac{\partial W}{\partial [E_1]} = \frac{\partial W}{\partial [E_2]}$

This condition straightforwardly requires the following quadratic equation to hold:

$$([E_1])^2 + K[E_1] - (2[E_2] + K)([E_2] + K) = 0,$$

which can be rewritten as:

$$[E_1] = \frac{-K + (K^2 + 4(2[E_2] + K)([E_2] + K))^{1/2}}{2} \tag{S7}$$

Finally, one can distinguish this optimal allocation depending on the level of saturation of the reactions:

$$\begin{cases} [E_1] \approx \sqrt{2}[E_2], \text{ if } [E_2] \gg K & \text{(S8a)} \\ [E_1] \approx 2[E_2], \text{ if } [E_2] \longrightarrow K & \text{(S8b)} \\ [E_1] \text{ is independent from } [E_2], \text{ if } K \gg [E_2] & \text{(S8c)} \end{cases}$$

As a conclusion, it means that a cell should allocate more to the first part of the pathway since it contributes more to the fitness (both directly and indirectly), a phenomenon which fades away, to a certain extent, when reactions stand far from saturation. More precisely, cells should allocate around 0.6 of their available proteome to the first part of the pathway under such circumstances. This estimate needs not be considered as a quantitative prediction (despite being close to findings with the more realistic pathway), yet, for it relies on an oversimplified definition of the influence of enzyme concentration on fluxes. It is all the more true since the expression used to describe it originally comes from a framework that does not capture realistically reactions approaching saturation [Bagheri-Chaichian et al., 2003] - but see [Yi and Dean, 2019] and [Labourel and Rajon, 2021] for other approaches showing similar saturating effects of enzyme concentrations.

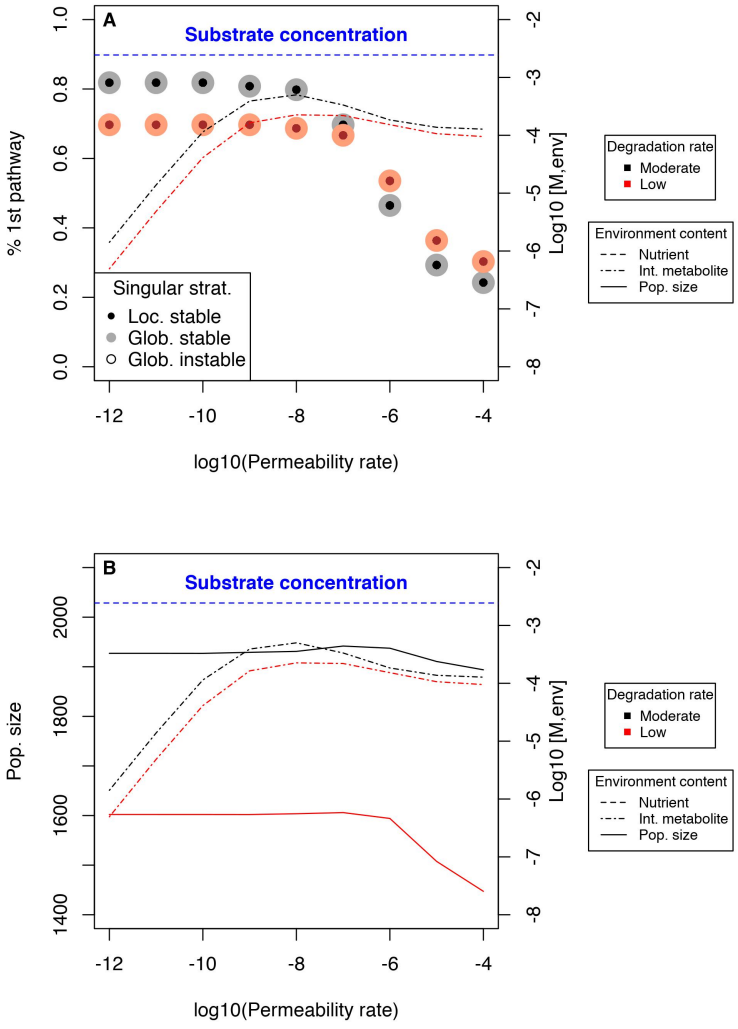

Figure S5: (A) shows the strategy that evolved when considering different degradation rates ($\eta = 10^{-3}s^{-1}$ and $\eta = 10^{-2}s^{-1}$) and enzymes highly efficient, for different permeability levels of the membrane - permeability only concerns one nutrient, in the middle of the chain. To cope with this phenomenon and avoid the cost of leakiness, cells allocate more to the second part of the pathway in either cases. Still, an ecological niche tends to emerge as the concentration of the intermediate metabolite in the environment increases steadily until moderate permeability levels are reached. They then slightly decrease as population sizes (B) also decrease. Remarkably, intermediate levels enable a slight increase in the population size, which means that the population is better at depleting its environment thanks to leakiness. No matter what, singular strategies are stable, which is explained by the large difference between the nutrient concentration in the environment (blue) and that of the intermediate metabolite.

Text S3 Membrane permeability and optimal allocation between pathways

Because fitness contributions add up along the pathway, we have shown that it may be relevant for an organism to favour certain of its reactions over others. Up until there, the selective pressures faced by enzymes were identical. Yet, the situation may turn out very differently, for instance if a metabolite is susceptible to be released in the environment, either passively through simple diffusion or actively through excretion machineries. In this section, we evaluated the impact on optimal metabolic strategies that membrane permeability may beget.

Text S3.1 High enzyme efficiencies foster cell investment in downstream enzymes

We first report what happens when enzyme kinetic efficiencies are high - approximately, one order of magnitude higher than median values observed in datasets [Bar-Even et al., 2011]. To grasp the qualitative effect of diffusion affecting a given metabolite, the contribution of each reaction to fitness, which depends on its flux, is first considered identical (each molecule produced by any reaction contribute one fitness unit). This echoes findings detailed in the article about the relevance for cells to cope with permeability by allocating more to the part of the pathway downstream the metabolite that is subject to it. One interesting phenomenon is that the higher population size is found for intermediate permeability levels, which means that permeability may help cells to deplete the environment more quickly - see Figure S5.

We have also reported in APPENDIX (Figure S13) the outcomes when accounting for metabolite toxicity, which proved to have very little impact.

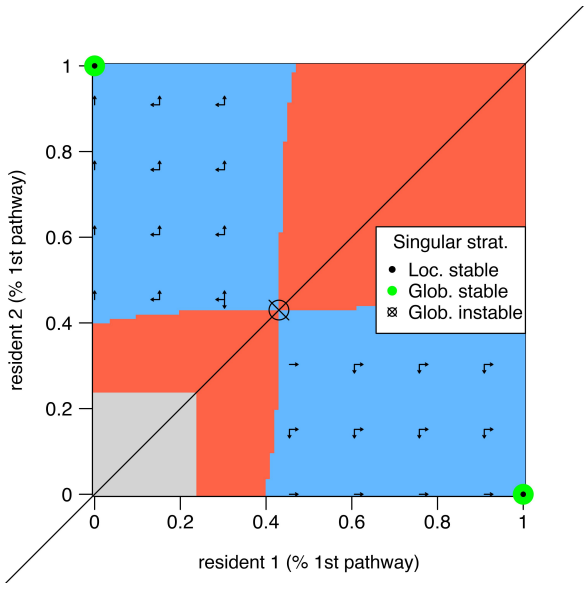

Figure S6: This TEP shows how a trait should evolve in a population comprised of two resident strategies. Each coalition is made up by two resident strategies, except on the bottom left toward upper right bisector for which the two resident strategies are identical. The red area is an area where coexistence is not possible, contrary to the blue one. In the blue area, we determine the fitness of each neighbouring mutant: there are four such mutants, except on the boundaries of the plot, as each resident can mutate and either increase or decrease its trait by one small unit. Here, we see that mutants that invade coalitions push the trait towards the upper left corner or the lower right one, which is exactly similar as these plots are symmetrical.

Text S3.2 Moderately high enzyme efficiencies foster cross-feeding diversification

In this section, kinetic parameters have been set to their median values found in Bar-Even et al. [2011]'s dataset, that is $k_f = 10^6 M^{-1} s^{-1}$ and $k_{cat} = k_r = 10^2 s^{-1}$. Results obtained with two degradation rates were reported in the article. We here provide details - see Figure S6 about the trait evolution plot (TEP) showing for which minimum permeability - combined with the moderate degradation rate - the singular strategy is invaded by a protected dimorphism [Geritz et al., 1998, Brännström et al., 2013]. Notice that owing to computational difficulty, it is not possible to decide whether the singular strategy is stable for cases where fitnesses are closer from one another. Typically, in this situation, one would expect that coalitions are favored where the subtypes are in between the singular strategies and the specialist ones [Geritz et al., 1998].

We also provide results in Figure S7 for two different degradation rates, which confirm observations described in the main body of the article.

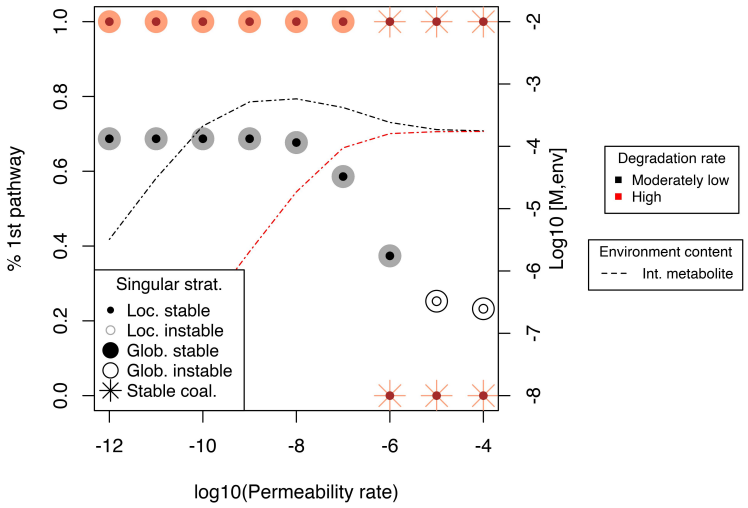

Figure S7: Outcomes of competition for two other degradaiton rates than those shown in the main document (moderately low=$10^{-2.5}s^{-1}$; high=$10^{-1.5}s^{-1}$). A moderately low degradation rate comes with the emergence of a new niche when permeability coefficients are high, but it remains dubious whether or not this should trigger diversification under the conditions studied here. On the contrary, a high degradation rate yields a high selective pressure and should therefore trigger diversification between two specialist phenotypes when the permeability P exceeds $10^{-6}dm \cdot s^{-1}$

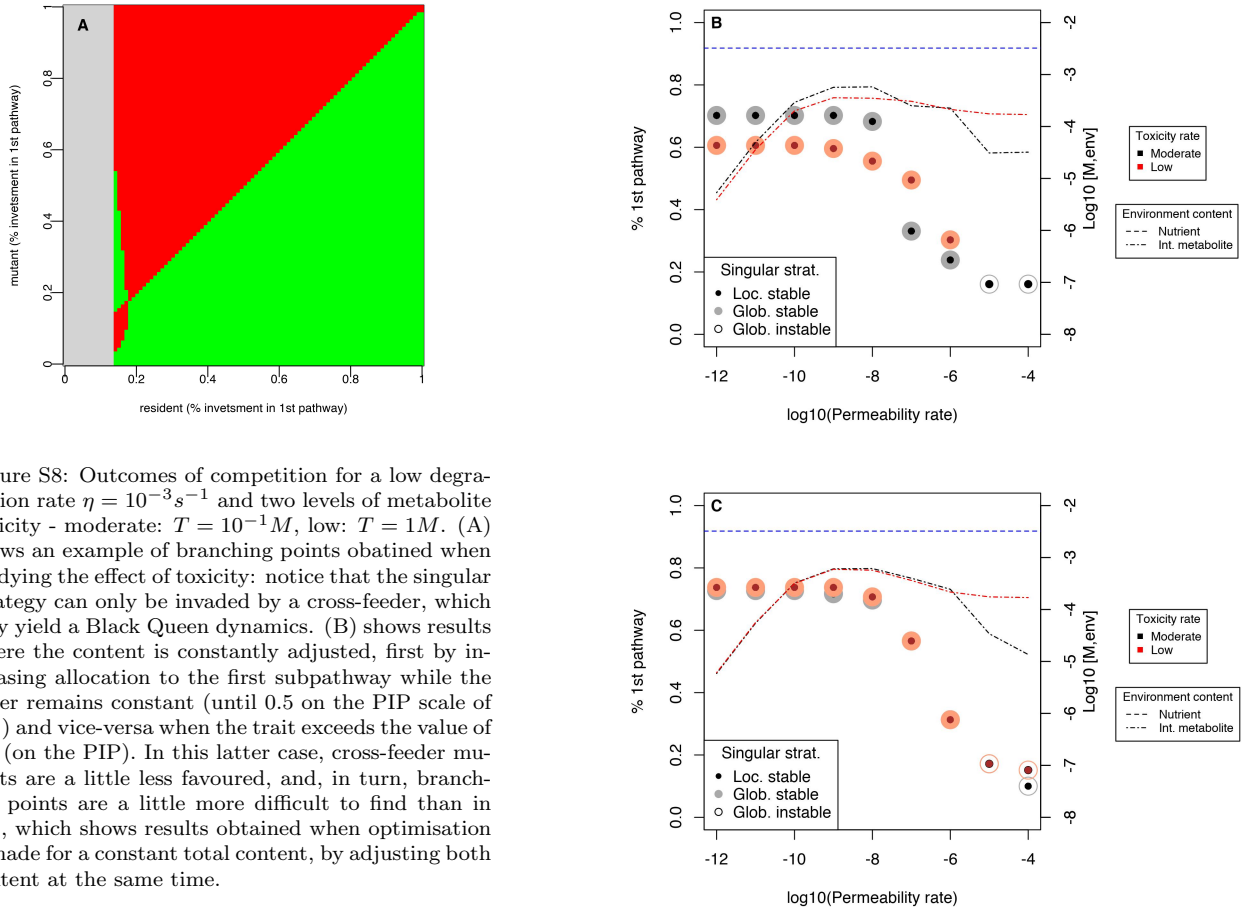

Figure S8: Outcomes of competition for a low degradation rate $\eta = 10^{-3}s^{-1}$ and two levels of metabolite toxicity - moderate: $T = 10^{-1}M$, low: $T = 1M$. (A) shows an example of branching points obatined when studying the effect of toxicity: notice that the singular strategy can only be invaded by a cross-feeder, which may yield a Black Queen dynamics. (B) shows results where the content is constantly adjusted, first by increasing allocation to the first subpathway while the other remains constant (until 0.5 on the PIP scale of (A)) and vice-versa when the trait exceeds the value of 0.5 (on the PIP). In this latter case, cross-feeder mutants are a little less favoured, and, in turn, branching points are a little more difficult to find than in (C), which shows results obtained when optimisation is made for a constant total content, by adjusting both content at the same time.

Text S3.3 Toxicity may also yield cross-feeding interactions

We tested how metabolite toxicity influences the outcomes by setting $\eta = 10^{-3}s^{-1}$ as to limit the impact of this parameter and studied two toxicity levels: $T = 10^{-1}M$ and $T = 1M$. The outcomes are qualitatively different although branching points still emerge (see Figure S8 - A). Because there is little loss of metabolites, the generalist strategy has no interest to sacrifice its second subpathway, which looks like a Black Queen coexistence where it may seem costly to keep the first subpathway but it is nonetheless essential for the

community to survive. Yet, for very high permeability rates, cross-feeders may eventually have the edge on generalist strategies (see Figure S8 - B and C), although the path leading to them is narrow, at least on a linear scale. This may not undermine these findings since levels of expression are likely evolving on a log scale, and switching off a subpathway may still be achievable through few regulatory mutations. Notice that for this latter reason, we did not draw coexistence plots because most of them would be dubious since weak mutations should not permit the invasion by cross-feeders; only simulations would be able to determine whether or not cross-feeding evolves and, if so, how often it does. Besides, we have also shown that these results are robust to the way mutations are drawn, as PIPs based on mutants that change only one of their concentration (and set the other to the optimum without permeation) display similar trends. This situation is slightly different because cells may be able to decrease their total content in order to avoid excess crowding and its deleterious effect.

Text S3.4 Realistic conditions including reversibility also promote a similar proteome allocation and the emergence of cross-feeding

I this subsection, we report results when considering a realistic set of constraints combining both reversibility, toxicity and degradation. Because what matters when reversibility enters into play is both the degradation rate and constraints set by the initiating transporter (whose reversibility rate is close to 1), the optimal allocation still requires to focus on the first sub-pathway, despite the pressure also set by toxicity. Notice that because reversibility is spread all along the pathway, when it neighbours 1, the optimal allocation quickly switches to an equal allocation between downstream and upstream enzymes, while a higher reversibility would lead to the need of enhancing the expression of downstream enzymes, as the reversibility which would set the higher constraint would be that within the cell rather than that of the trasporter initiating the pathway. Notice that the average reversibility of the central carbon metabolism is approximately 10^{-1} leading to a very high departure from the equal spread. Assuming such a reversibility and different payoffs for each subpathway does not change qualitatively this finding, with an optimal allocation between 2:1 and 3:1.

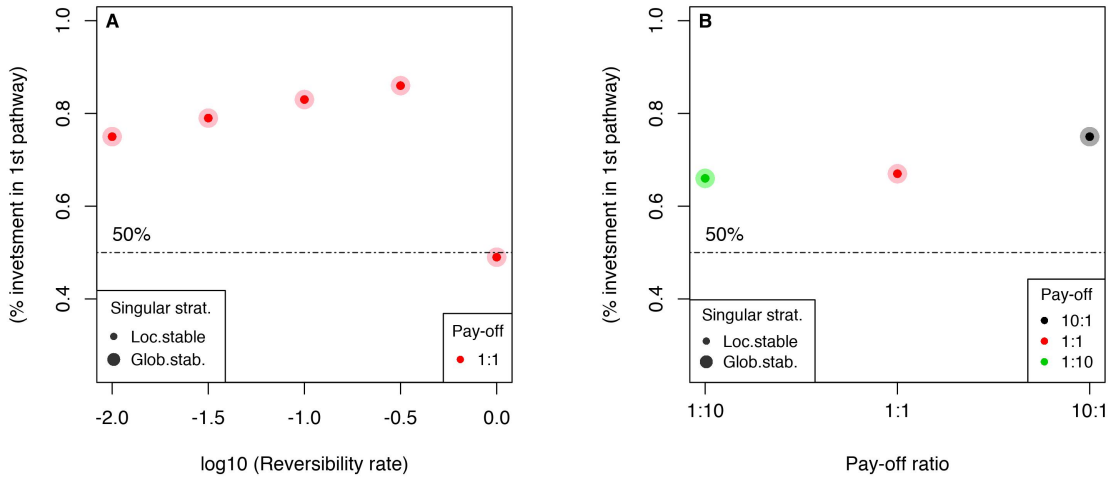

Figure S9: Influence of reversibility on the optimal allocation. In (A), the influence of reaction reversibility is shown for an equal yield of each subpathway when a low toxicity ($T = 1M$) and a low degradation rate ($\eta_d = 10^{-4}s^{-1}$) are considered. In (B), reversibility is set to its geometric mean for enzymes involved in central carbon metabolism while toxicity ($T = 10^{-1}M$) and the degradation rate ($\eta = 10^{-3}s^{-1}$) are set to moderate values and the ratio between sub-pathways yields is proven to be of little influence.

Finally, we report thereafter the influence of the membrane leakage when assuming realistic internal and metabolic constraints. In these cases again, cross-feeding interactions only arise when permeability reaches very high values $P > 10^{-6}dm \cdot s^{-1}$, and, as in the toxicity case, it may involve a black queen dynamics where the stable coalition is composed of a generalist strategy focusing on the second part of the pathway and a specialist exploiting only the second sub-pathway. Notice that when incorporating reversibility in the model, it is not easy to find an algorithm that allows to find the demographic equilibrium of coalitions with sufficient precision; thus, we do not report these dubious results here though they tend to point towards the emergence of coalitions.

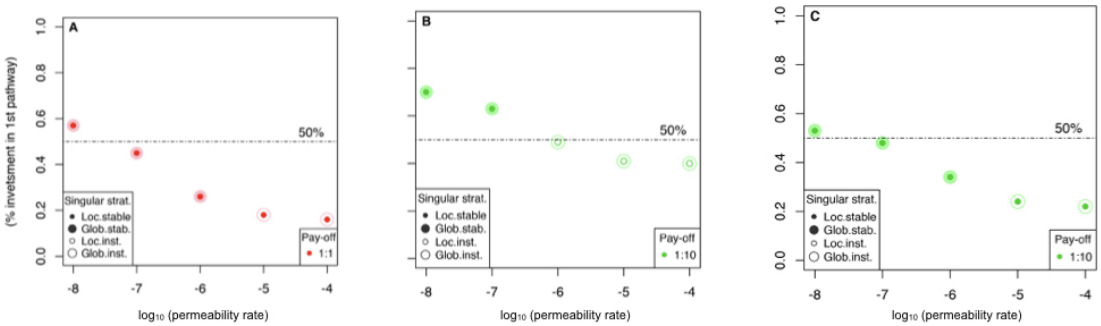

Figure S10: Influence of the permeability rate when assuming realistic conditions for toxicity, reversibility and degradation - $\eta_d = 10^{-3}s^{-1}$, $T = 10^{-1}M$ and $K_{rev} = 10^{-1}$ in (A). In (B), results coincide with a high degradation rate ($\eta_d = 10^{-2}s^{-1}$) while in (C), we report results when a higher toxicity rate $T = 10{-}2M$ is considered. Cross-feeding interactions should generally be favored by evolution when permeability is high (although it is not possible to solve the TEPs when reversibility is also involved).

References

Steven S Andrews. Effects of surfaces and macromolecular crowding on bimolecular reaction rates. *Phys Biol*, 17(4):045001, May 2020. ISSN 1478-3975 (Electronic); 1478-3967 (Linking). doi: 10.1088/1478-3975/ab7f51.

Homayoun Bagheri-Chaichian, Joachim Hermisson, Juozas R. Vaisnys, and Günter P. Wagner. Effects of epistasis on phenotypic robustness in metabolic pathways. *Mathematical Biosciences*, 184(1):27–51, 2003. doi: https://doi.org/10.1016/S0025-5564(03)00057-9. URL https://www.sciencedirect.com/science/article/pii/S0025556403000579.

Arren Bar-Even, Elad Noor, Yonatan Savir, Wolfram Liebermeister, Dan Davidi, Dan S. Tawfik, and Ron Milo. The moderately efficient enzyme: Evolutionary and physicochemical trends shaping enzyme parameters. *Biochemistry*, 50(21):4402–4410, 05 2011. doi: 10.1021/bi2002289. URL https://doi.org/10.1021/bi2002289.

Pablo M. Blanco, Josep Lluís Garcés, Sergio Madurga, and Francesc Mas. Macromolecular diffusion in crowded media beyond the hard-sphere model. *Soft Matter*, 14:3105–3114, 2018. doi: 10.1039/C8SM00201K. URL http://dx.doi.org/10.1039/C8SM00201K.

Evert Bosdriesz, Meike T. Wortel, Jurgen R. Haanstra, Marijke J. Wagner, Pilar de la Torre Cortés, and Bas Teusink. Low affinity uniporter carrier proteins can increase net substrate uptake rate by reducing efflux. *Scientific Reports*, 8(1):5576, 2018. doi: 10.1038/s41598-018-23528-7. URL https://doi.org/10.1038/s41598-018-23528-7.

Åke Brännström, Jacob Johansson, and Niels Von Festenberg. The hitchhiker's guide to adaptive dynamics. *Games*, 4(3):304–328, 2013. ISSN 2073-4336. doi: 10.3390/g4030304. URL https://www.mdpi.com/2073-4336/4/3/304.

G E Briggs and J B Haldane. A note on the kinetics of enzyme action. *The Biochemical journal*, 19(2): 338–339, 1925. doi: 10.1042/bj0190338. URL https://pubmed.ncbi.nlm.nih.gov/16743508.

Hsin-Hung Chou, Nigel F. Delaney, Jeremy A. Draghi, and Christopher J. Marx. Mapping the fitness landscape of gene expression uncovers the cause of antagonism and sign epistasis between adaptive mutations. *PLOS Genetics*, 10(2):e1004149–, 02 2014. URL https://doi.org/10.1371/journal.pgen.1004149.

Ken A. Dill, Kingshuk Ghosh, and Jeremy D. Schmit. Physical limits of cells and proteomes. *Proceedings of the National Academy of Sciences*, 108(44):17876, 11 2011. doi: 10.1073/pnas.1114477108. URL http://www.pnas.org/content/108/44/17876.abstract.

R. John Ellis. Macromolecular crowding: obvious but underappreciated. *Trends in Biochemical Sciences*, 26 (10):597–604, 2020/03/20 2001. doi: 10.1016/S0968-0004(01)01938-7. URL https://doi.org/10.1016/S0968-0004(01)01938-7.

S. A. H. Geritz, E. Kisdi, G. Meszena, and J. A. J. Metz. Evolutionarily singular strategies and the adaptive growth and branching of the evolutionary tree. *Evolutionary Ecology*, 12(1):35–57, 1998. doi: 10.1023/A:1006554906681. URL https://doi.org/10.1023/A:1006554906681.

D L Hartl, D E Dykhuizen, and A M Dean. Limits of adaptation: the evolution of selective neutrality. *Genetics*, 111(3):655–674, 11 1985. URL https://pubmed.ncbi.nlm.nih.gov/3932127.

R Heinrich and T A Rapoport. A linear steady-state treatment of enzymatic chains. general properties, control and effector strength. *Eur J Biochem*, 42(1):89–95, Feb 1974. ISSN 0014-2956 (Print); 0014-2956 (Linking). doi: 10.1111/j.1432-1033.1974.tb03318.x.

R Heinrich, S Schuster, and H G Holzhütter. Mathematical analysis of enzymic reaction systems using optimization principles. *Eur J Biochem*, 201(1):1–21, Oct 1991. ISSN 0014-2956 (Print); 0014-2956 (Linking). doi: 10.1111/j.1432-1033.1991.tb16251.x.

H Kacser and J A Burns. The control of flux. *Symp Soc Exp Biol*, 27:65–104, 1973. ISSN 0081-1386 (Print); 0081-1386 (Linking).

Moshe Kafri, Eyal Metzl-Raz, Ghil Jona, and Naama Barkai. The cost of protein production. *Cell Reports*, 14(1):22–31, 2016. doi: https://doi.org/10.1016/j.celrep.2015.12.015. URL https://www.sciencedirect.com/science/article/pii/S221112471501428X.

Miriam Kaltenbach and Nobuhiko Tokuriki. Dynamics and constraints of enzyme evolution. *J Exp Zool B Mol Dev Evol*, 322(7):468–487, Nov 2014. ISSN 1552-5015 (Electronic); 1552-5007 (Linking). doi: 10.1002/jez.b.22562.

Edda Klipp and Reinhart Heinrich. Evolutionary optimization of enzyme kinetic parameters; effect of constraints. *Journal of Theoretical Biology*, 171(3):309–323, 1994. doi: https://doi.org/10.1006/jtbi.1994.1234. URL http://www.sciencedirect.com/science/article/pii/S0022519384712343.

Stefan Klumpp, Matthew Scott, Steen Pedersen, and Terence Hwa. Molecular crowding limits translation and cell growth. *Proceedings of the National Academy of Sciences*, 110(42):16754, 10 2013. doi: 10.1073/pnas.1310377110. URL http://www.pnas.org/content/110/42/16754.abstract.

Florian Labourel and Etienne Rajon. Resource uptake and the evolution of moderately efficient enzymes. *Molecular Biology and Evolution*, 5/10/2021 2021. doi: 10.1093/molbev/msab132. URL https://doi.org/10.1093/molbev/msab132.

Xin Li, Fan Wu, Feng Qi, and Daniel A. Beard. A database of thermodynamic properties of the reactions of glycolysis, the tricarboxylic acid cycle, and the pentose phosphate pathway. *Database*, 2011, 9/27/2021 2011. doi: 10.1093/database/bar005. URL https://doi.org/10.1093/database/bar005.

Michael Lynch and Georgi K Marinov. The bioenergetic costs of a gene. *Proceedings of the National Academy of Sciences of the United States of America*, 112(51):15690–15695, 12 2015. doi: 10.1073/pnas.1514974112. URL http://www.ncbi.nlm.nih.gov/pmc/articles/PMC4697398/.

Andreas Maier, Bernhard Völker, Eckhard Boles, and Günter Fred Fuhrmann. 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, 12 2002. ISSN 1567-1356. doi: 10.1111/j.1567-1364.2002.tb00121.x. URL https://doi.org/10.1111/j.1567-1364.2002.tb00121.x.

Benno H. ter Kuile and Michael Cook. The kinetics of facilitated diffusion followed by enzymatic conversion of the substrate. *Biochimica et Biophysica Acta (BBA) - Biomembranes*, 1193(2):235–239, 1994. doi: https://doi.org/10.1016/0005-2736(94)90158-9. URL http://www.sciencedirect.com/science/article/pii/0005273694901589.

B Teusink, J A Diderich, H V Westerhoff, K van Dam, and M C Walsh. 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, 02 1998. URL https://pubmed.ncbi.nlm.nih.gov/9457857.

Andreas Wagner. Energy Constraints on the Evolution of Gene Expression. *Molecular Biology and Evolution*, 22(6):1365–1374, 03 2005. ISSN 0737-4038. doi: 10.1093/molbev/msi126. URL https://doi.org/10.1093/molbev/msi126.

Xiao Yi and Antony M Dean. Adaptive landscapes in the age of synthetic biology. *Molecular biology and evolution*, 36(5):890–907, 05 2019. doi: 10.1093/molbev/msz004. URL https://pubmed.ncbi.nlm.nih.gov/30657938.

APPENDIX - Subset of PIPs obtained for the optimal pathway content

We here report PIPs that were obtained when testing the influence of enzyme efficiencies and the concentration of the first enzyme:

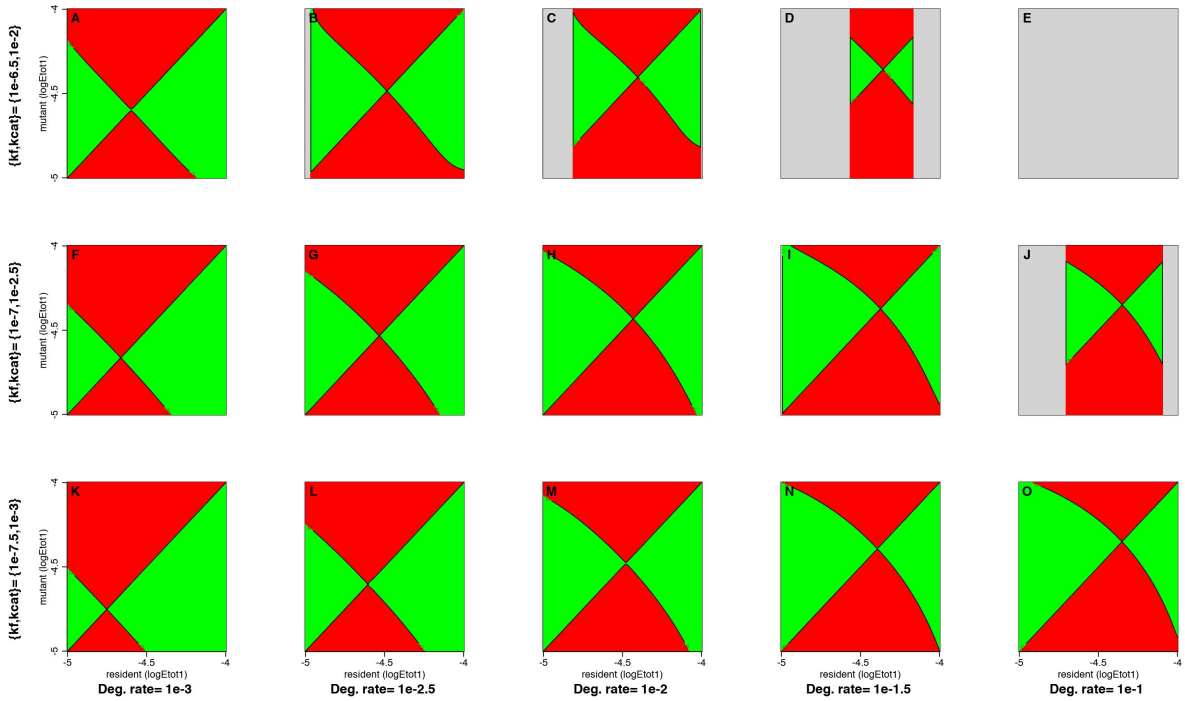

Figure S11: PIPs showing the convergent stable strategy - CSS hereafter - for a first enzyme concentration (the enzyme directly following the transporter) $[E_{tot,0}] = 100\mu M$,different average enzyme efficiencies (rows) and degradation rates (columns). The point at the crossroads of green and red areas denote the CSS. Grey areas represent areas where no strategy at all is viable, which means that no strategy is able to produce enough energy to compensate both for the protein production cost and the minimumx flux ensuring that births compensate for deaths. Results are shown and commented more in depth on figure S1, where CSSs are denoted as points.

We also report PIPs that were used to determine how the content between two parts of pathways should establish with low selective constraints and how it should change in response to higher selective ones:

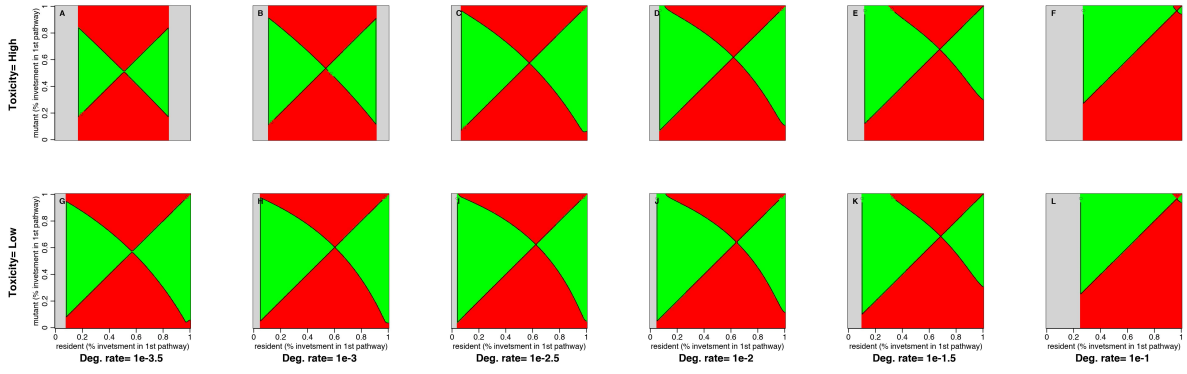

Figure S12: PIPs showing the convergent stable strategy for the set of generic parameters in the situation where an organism can spread its proteome between upstream reactions and downstream ones. Axis denote the resident and the mutant strategy, which are expressed as the investment in the first part of the pathway as the total concentration is set to its optimum (found in section Text-S1). Results in Figure 1 of the article sums up these PIPs by showing only the CSSs.

We report below the outcomes about optimal allocation obtained when considering toxicity, a low degradation rate and high enzyme efficiencies:

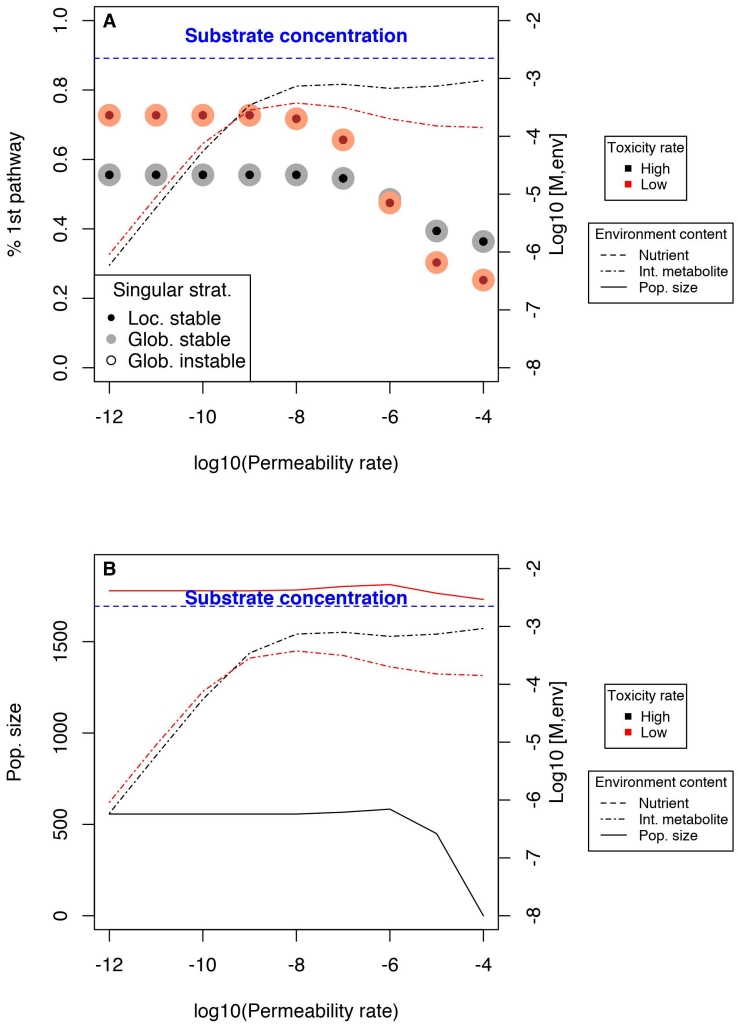

Figure S13: (A) shows the strategy that evolved when considering different a low degradation rate ($\eta = 1e-3s^{-1}$) and various toxicity rates (High: $T = 1e-2M$ and Low: $T = 1e-1M$) and enzymes highly efficient, for different permeability levels of the membrane - permeability only concerns one nutrient, in the middle of the chain. Results are very similar than those obtained when we did not consider toxicity, except that the population crashes with high toxicity and permeability rates.

We report the underlying TEPs that helped draw Figure S7:

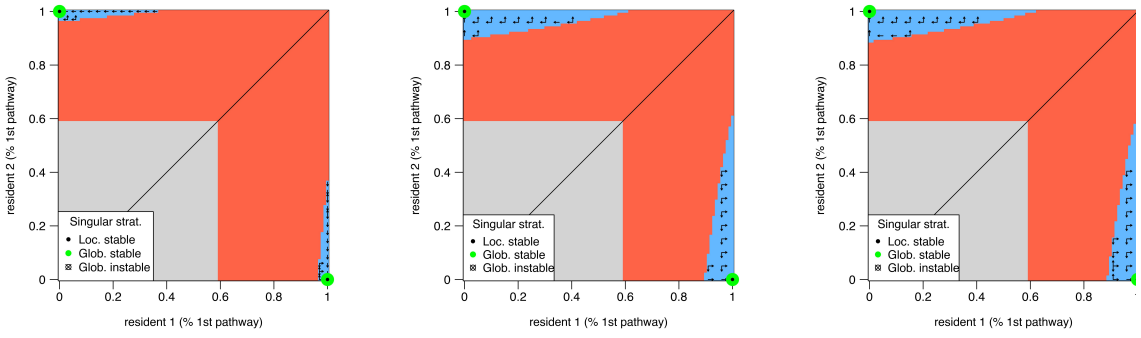

Figure S14: TEPs in the case of median enzyme efficiencies ($k_f = 10^6 M^{-1}s^{-1}$, $k_{cat} = 10^2 s^{-1}$) and a high degradation rate $\eta = 10^{-2}s^{-1}$. Evolution yields the same strategy in all these cases, that is the coexistence of two specialist strategies where the second subtype cross-feeds on the first one.

We also report the TEP which helped draw the second figure of the article, when considering a degradation rate $\eta_d = 10^{-5}s^{-1}$:

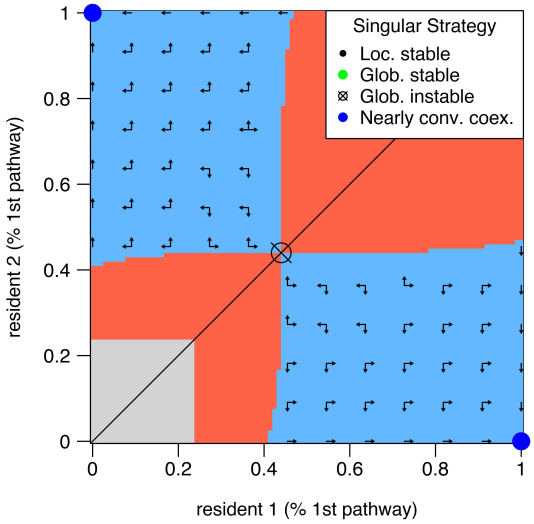

Figure S15: This TEP shows the results when considering only a degradation rate, when the permeability rate is set to $P = 10^{-5} dm\ s^{-1}$, as in section Text S3.2

Finally, we report the results which helped draw PIPs when realistic conditions were considered, that is to say including reactions reversibility besides metabolite toxicity and a relatively low degradation rate. These sets of PIPs were used to draw Figure S10 of SM, where plos on the left coincide with low permeability rates while plots on the right coincide with the highest permeability rates: