J Biol Phys (2008) 34:73–90
DOI 10.1007/s10867-008-9067-2

ORIGINAL PAPER

A Strategy to Calculate the Patterns of Nutrient Consumption by Microorganisms Applying a Two-Level Optimisation Principle to Reconstructed Metabolic Networks

Miguel Ponce de León · Héctor Cancela · Luis Acerenza

Received: 30 October 2007 / Accepted: 19 March 2008 /
Published online: 14 May 2008

Abstract Bacterial responses to environmental changes rely on a complex network of biochemical reactions. The properties of the metabolic network determining these responses can be divided into two groups: the stoichiometric properties, given by the stoichiometry matrix, and the kinetic/thermodynamic properties, given by the rate equations of the reaction steps. The stoichiometry matrix represents the maximal metabolic capabilities of the organism, and the regulatory mechanisms based on the rate laws could be considered as being responsible for the administration of these capabilities. Post-genomic reconstruction of metabolic networks provides us with the stoichiometry matrix of particular strains of microorganisms, but the kinetic aspects of in vivo rate laws are still largely unknown. Therefore, the validity of predictions of cellular responses requiring detailed knowledge of the rate equations is difficult to assert. In this paper, we show that by applying optimisation criteria to the core stoichiometric network of the metabolism of *Escherichia coli*, and including information about reversibility/irreversibility only of the reaction steps, it is possible to calculate bacterial responses to growth media with different amounts of glucose and galactose. The target was the minimisation of the number of active reactions (subject to attaining a growth rate higher than a lower limit) and subsequent maximisation of the growth rate (subject to the number of active reactions being equal to the minimum previously calculated). Using this two-level target, we were able to obtain by calculation four fundamental behaviours found experimentally: inhibition of respiration at high glucose concentrations in aerobic conditions, turning on of respiration when glucose decreases, induction of galactose utilisation when the system is depleted of glucose and simultaneous use of glucose and galactose as carbon sources when both sugars are present in low concentrations. Preliminary results of the coarse pattern of sugar utilisation were also obtained with a genome-scale *E. coli* reconstructed network, yielding similar qualitative results.

M. Ponce de León · L. Acerenza (✉)
Laboratorio de Biología de Sistemas, Facultad de Ciencias, Universidad de la República, Iguá 4225, Montevideo 11400, Uruguay
e-mail: aceren@fcien.edu.uy

H. Cancela
Instituto de Computación, Facultad de Ingeniería, Universidad de la República, Montevideo, Uruguay

Keywords Metabolic networks · Metabolic responses · Metabolic regulation · Genome-scale networks · Stoichiometric networks · Optimisation criteria · Metabolic models · *Escherichia coli* metabolism

1 Introduction

With the advent of the post-genomic era, it is possible to reconstruct the metabolic map of particular strains of microorganisms. These stoichiometric networks include all the metabolic processes that could take place in the cell under any particular experimental conditions. They provide genome-scale information about the maximal metabolic capabilities of the organisms, compatible with their stoichiometric constraints. Metabolic functioning is also subject to dynamic and regulatory constraints imposed by the types of rate laws that govern the kinetics and thermodynamics of the reaction steps. However, genome-scale knowledge about the kinetic aspects of in vivo rate laws is still not available.

The profuse regulatory mechanisms operating in the living cell could be seen as the way to manage maximal metabolic capabilities to respond in appropriate ways to the challenges imposed by diverse environmental conditions. For example, microorganisms show genetic regulatory adaptations when the nutrients in the medium change. In particular, when *Escherichia coli* is grown in mixtures of two sugars, e.g. glucose and galactose, the pattern of consumption depends on the amount and proportions of the sugars. In a medium abundant in glucose and galactose, the bacteria start growing using glucose only. Once glucose is exhausted, the enzymes required for galactose utilisation are genetically induced, resulting in biphasic behaviour [1]. In contrast, studies of bacterial growth in mixtures of glucose and galactose, both at low concentrations, show that the sugars are consumed simultaneously [2–4].

The survival of an organism strongly depends on its ability to respond to environmental changes in appropriate ways, e.g. avoiding or neutralising harmful situations and exploiting advantageous ones. These adaptive responses rely on regulatory mechanisms that are the result of a long evolutionary time course during which the responses of metabolism to environmental changes have been improved and, perhaps, optimised. The overall goal of this putative optimisation process is, simply, the survival of the organism. However, its mathematical formulation, i.e. the target function expressed in terms of system variables, may depend on the type of response and is usually difficult to grasp. In addition, if two competing organisms have evolved to respond equally well to changes in the environment, the one that shows the higher growth rate survives [5]. Therefore, "the survival of the fittest" implies that the fastest of those organisms presenting the optimal responses for adaptation to the environment survives [6].

Predicting metabolic responses has been a difficult task, which is due to the fact that the metabolic networks of even the simplest living organisms are highly complex. This complexity includes a high number of component metabolic intermediates, interconverted by an intricate network of biochemical reactions, whose rate laws are strongly non-linear. Moreover, as was mentioned above, the specific functions describing the rate laws operating in vivo are largely unknown. Hence, it is difficult to assert the validity of predictions about metabolic responses, obtained with differential equations models built with these rate equations. On the other hand, the stoichiometric network is the most extensively known information of metabolic systems. Models that rely mainly on this type of experimental data to make predictions of, for example, viability of mutants, have been highly successful (see, e.g. [7]). Therefore, one would expect that strategies designed to calculate metabolic responses based on the stoichiometric network would also lead to reliable predictions.

One way to calculate metabolic responses using the stoichiometric network is to apply optimisation criteria. For example, one could test if the responses experimentally shown when *E. coli* is grown in different mixtures of sugars may be obtained by applying optimisation criteria to its reconstructed metabolism. In this context, the question is if there exists one optimisation criterion that, applied to the maximal metabolic capabilities, allows us to reproduce the rich repertoire of responses exhibited by the bacteria when the nutrient composition of the medium is changed. This is the type of question that we aim to address in the present contribution.

2 The Strategy

Here, we shall develop a strategy to calculate the responses of *E. coli* when it is grown in mixtures of sugars present in different amounts and proportions. This will consist in applying an optimisation procedure to the metabolic network of the organism. For this purpose, we shall define the stoichiometric network, the constraints that bind the decision variables, and the optimisation target.

The stoichiometric network that we shall use is the so-called core network of *E. coli* [8] consisting on $m=49$ metabolites (nodes) and $r=64$ reactions (hyper-edges). This network represents, basically, a subset of the metabolic machinery used for glucose degradation. We have added five steps to this network to include the pathway of galactose incorporation. Therefore, the network used, represented in Fig. 1, has $m=54$ and $r=69$. The list of reactions is given in Appendix 1. In this list, in addition to the 69 internal reactions, there are 15 external reactions (whose names start by "EX_") representing auxiliary fluxes introduced to maintain the external pools constant so as to ensure that the system reaches a non-trivial steady state. As these are not real biochemical reactions, they are not included when the number of steps is calculated. The network can also be described by a stoichiometry matrix $\boldsymbol{N}$, the element n_{ij} being the stoichiometric coefficient of metabolite i in reaction j. In this matrix, both internal and external fluxes are included.

It is assumed that the rates are subject to two different types of constraints. The system is at a stable steady state. In this state, the rates are balanced, i.e. for every metabolite, the sum of the rates of production equals the sum of the rates of consumption. This condition is expressed mathematically as: $\boldsymbol{N.v}=\boldsymbol{o}$, where $\boldsymbol{N}$ is the stoichiometry matrix, $\boldsymbol{v}$ is the vector of rates and $\boldsymbol{o}$ is the zero vector. In addition, there are lower and upper limits to the values that the rates can take. For example, rates of nutrient incorporation from outside the cell are limited by the maximal rate of transport through the plasma membrane that could be achieved and rates of irreversible reactions are positive, their lower limit being zero. In Appendix 1, the lower and upper limits of the external fluxes are listed. Note that all calculations are performed in aerobic conditions, which is implemented allowing an upper limit of the external oxygen flux equal to 10.

Finally, the target of optimisation is defined. Several optimisation targets have been proposed for the evolution of metabolic systems. The variables associated with these targets include fluxes, metabolite concentrations, response time, productivity, net ATP production and number of reaction steps, among others ([9–15] and references therein). Here, we shall assume that, in the evolution of cellular responses shown when the nutrients in the medium change, there are two optimisation problems involved. The main target of evolution (i.e. the master problem) is the maximisation of growth rate, and the secondary target (i.e. the lower level problem) is the minimisation of the number of active reaction steps. Mathematically, the lower level problem functions as a constraint that has to be solved before the master

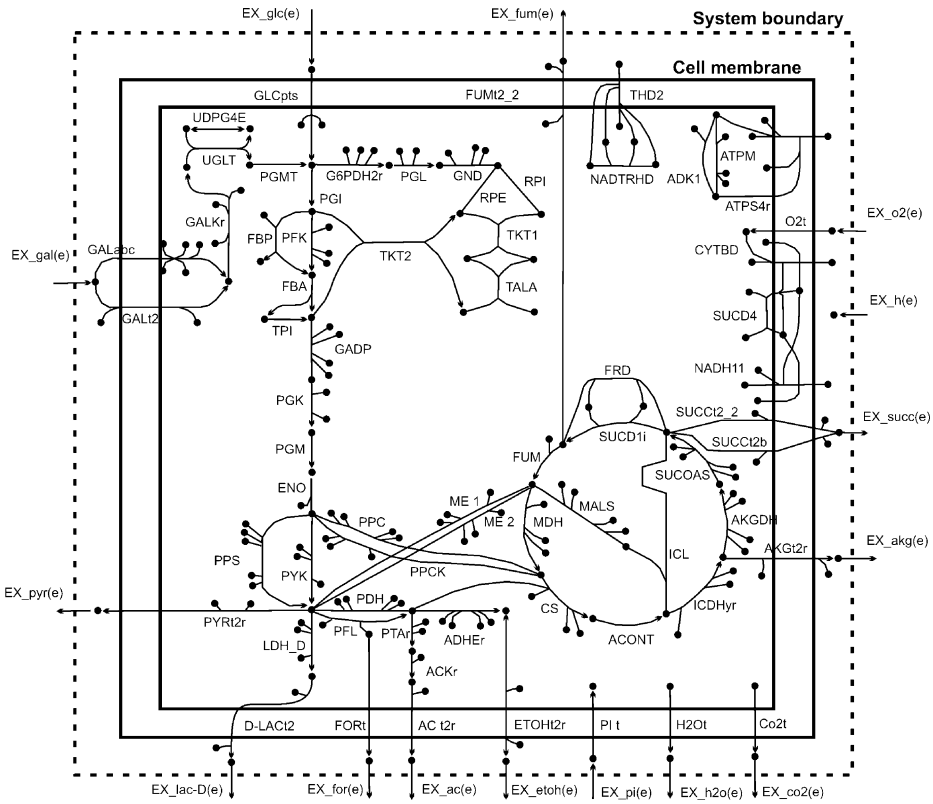

Fig. 1 Diagrammatic representation of the core metabolic network of *E. coli*. The reaction corresponding to v_{growth}, defined in Appendix 1, is not represented (adapted from [8])

problem. In this particular case, the two problems are separable and can be successively solved. Firstly, the lower level problem is solved. Implementation of this optimisation procedure requires introducing r integer variables, y_j, which can take only two values: 0 (reaction step off) and 1 (reaction step on). The number of active reaction steps is calculated as $n = \sum y_j$ (see model in Appendix 2). The minimum of n, $n_{\min}$, is obtained subject to the condition that the growth rate, v_{growth} (defined in Appendix 1), is higher than a certain lower limit, v^1_{growth}. Secondly, the master problem is solved. The maximum of v_{growth} is obtained subject to the condition that the number of active reaction steps is equal to the minimum previously calculated. This two-stage process in which, first, n is minimised (conditioned to $v_{\text{growth}} \geq v^1_{\text{growth}}$) and, second, v_{growth} is maximised (subject to $n{=}n_{\min}$) is a bi-level mixed integer linear programming optimisation problem. It is important to note that optimisation according to the secondary target alone may result in alternate flux distribution optima with the same minimum number of active reaction steps. When the main target is applied, normally, one flux distribution having the maximum growth rate achievable with the minimum number of steps is selected. Two reasons justify the use of this two-level target of optimisation. First, a plausible explanation can be elaborated arguing why the minimisation of the number of active reaction steps could lead to an increase in growth rate (see Section 3). Second, some of the theoretical predictions obtained coincide with fundamental experimental findings (see Section 4). The calculations are performed (using

the optimisation software CPLEX 10.1) for different combinations of values of the upper limits that the input rates of glucose $\left(v^{\mathrm{u}}_{\mathrm{Glc}}\right)$ and galactose $\left(v^{\mathrm{u}}_{\mathrm{Gal}}\right)$ can take. v^{u}_{Glc} and v^{u}_{Gal} were between 0 and 20, taking all combinations of the values at intervals of 0.1 (40,000 points). This number of points is sufficiently large to identify the boundaries between regions of different behaviours. For each pair of values of v^{u}_{Glc} and v^{u}_{Gal}, the overall procedure would give the "fittest organism" which, in our context, is the fastest growing organism that satisfies the condition of minimum number of active steps required for optimal response.

3 Number of Active Reaction Steps as a Target of Optimisation

In the previous section, we proposed that, in the evolution of the cellular responses to changes in the nutrients of the medium, the target was the minimisation of the number of active reaction steps (subject to v_{growth} greater than a certain lower limit) and subsequent maximisation of v_{growth} (keeping the number of steps at the minimum found). It is easy to understand why the maximisation of growth rate has been a central target in evolution. As we have discussed above, from all the organisms that respond in the appropriate ways, the fastest is the one that survives. In contrast, why the minimisation of the number of active reaction steps has been an important constraint in evolution requires a more elaborated justification.

The main argument that we will develop is that the minimisation of the number of steps indirectly contributes to increase growth rate. There are two ways in which this occurs: "nutrient reallocation" and "protein reallocation". The first is that reducing the number of steps that consume nutrient for other purposes than growth, the nutrient saved can be used to increase growth rate. The second is that reducing the number of steps that are not indispensable for growth, the protein used in these processes can be reallocated to the active growth-producing processes, increasing growth rate.

Let us first discuss how reducing the number of active reaction steps can lead to nutrient reallocation and, consequently, to the increase in growth rate. We assume that glucose is the only carbon source, its input being fixed. If we turn off all reaction steps that do not contribute to growth, then glucose that was consumed in these processes can be used to produce biomass. Therefore, the minimisation of the number of active steps (subject to attaining a non-zero growth) would give an active sub-network where the subsequent maximisation of the growth rate may result in a high growth-rate value.

There is experimental evidence, that during evolution, decrease in the number of active reaction steps and increase in growth rate can occur in parallel. For instance, experimental results show that, during *E. coli* evolution under controlled conditions, fitness increases, while unused catabolic functions decay, this decay being beneficial and therefore fixed by selection [16]. A model developed to explain these and other important features of *E. coli* evolution is precisely based on the assumption that due to the decay of unused catabolic functions, nutrient that was previously used for adaptation is reallocated to produce growth [17, 18].

Next we discuss how reducing the number of active reaction steps can lead to protein reallocation and, as a consequence, to the increase in growth rate.

If we turn off all reaction steps that do not contribute to growth, then the amount of protein catalysing these steps could be reallocated to catalyse steps that produce biomass. Therefore, the minimisation of the number of active steps (subject to attaining a non-zero growth) would give an active sub-network where each step has a higher amount of enzyme, resulting in a higher growth rate value.

In addition, the amount of protein that can be accommodated in the volume of the cell has an upper limit. A direct consequence of this fact is that shorter pathways can carry higher fluxes than longer ones, all other properties being equal [14]. For instance, let us consider the four step pathway (pathway I) represented in Fig. 2a. In this pathway, X and P are fixed by external means; they are parameters, while S_i, i=1, 2, 3, are variables. We compare this pathway with a shorter one (pathway II) obtained from pathway I by removing steps 1 and 2, and keeping S_2 at a fixed value X_2, equal to the steady state value of S_2 in pathway I (Fig. 2b). Pathway II shows the same steady-state flux as pathway I, but this flux is obtained with less amount of enzyme. The protein that catalysed the two steps that have been removed (1 and 2) can be reallocated to steps 3 and 4 of pathway II. If we assume that the rates of steps 3 and 4 are proportional to the corresponding enzyme concentrations and that reallocation is done such that both enzyme concentrations are increased by the same factor, the flux will also be increased by that factor [19, 20]. In summary, an increase in the flux can be obtained eliminating a portion of an unbranched chain of reactions and reallocating the protein to the remaining steps (what keeps the total amount of protein fixed).

The branched pathway represented in Fig. 2c has two different input limbs, with one and two steps, respectively. We assume that these two limbs have the same amount of protein and that steps 1 and 2 are governed by the same rate laws, with the concentrations of S_1 and X_1 being equal. Under these conditions, the shorter pathway will produce a higher flux (see above). Therefore, if we eliminate the input limb with the highest number of steps and we adequately reallocate the corresponding enzyme concentration, the output flux to P will increase. Note that this type of scheme is similar (qualitatively) to the first reaction steps through which glucose and galactose enter the metabolism (see Fig. 1).

In summary, the minimisation of the number of active reaction steps followed by protein reallocation (maintaining total protein concentration fixed) can be performed in such a way that growth rate increases. Moreover, this process tends to give a high flux per step. This is a desirable property, as the higher the fluxes the lower the deleterious effects due to fluctuations, which appear to be an inherent characteristic of living cells (see for example [21]).

Reduction in the number of active reaction steps in the metabolic network is performed by the cell turning off the expression of the genes coding for the corresponding enzyme activities. As a consequence, production of the enzyme protein ceases, but the existing enzyme remains active. In the next section, we shall report the results of minimising the

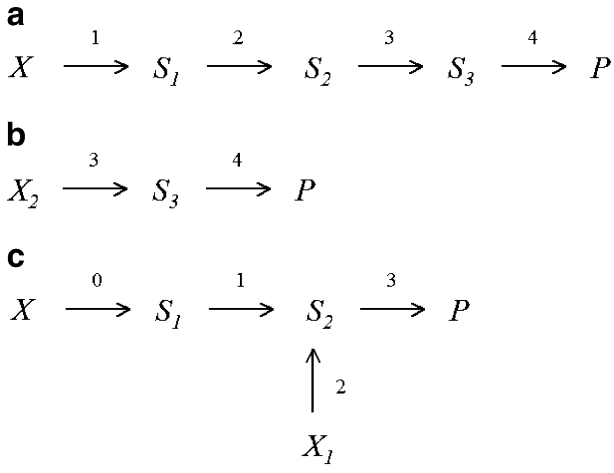

Fig. 2 Metabolic pathways: **a** unbranched four step pathway, **b** unbranched two step pathway and **c** branched pathway

number of active reaction steps (and subsequent maximisation of growth rate) in the core metabolic network of *E. coli*. In the solutions obtained, the fact that an enzyme catalysed step is turned off indicates that the gene corresponding to that enzyme is not expressed. However, this does not mean that the activity of the protein enzyme is absent in the cell if in the immediately previous experimental conditions the gene was on. Therefore, the results obtained, when moving from one experimental condition to another, should be interpreted, in the short run, more as the change in a metabolic flux produced by the change in the external conditions than the resulting flux in the final conditions.

4 Results

The two-level optimisation procedure described above (see Section 2) was applied to the "core" of *E. coli*, obtaining the following results.

In Fig. 3, we show the pattern of glucose and galactose consumption as a function of the pairs (v^u_{Glc}, v^u_{Gal}), i.e. the upper limits of the input rates of glucose and galactose. In the black region, there is no solution to the problem because, with the low values of v^u_{Glc} and v^u_{Gal} available, the condition that the growth rate should be greater than a lower limit $\left(v_{growth} \geq v^l_{growth} = 0.1\right)$ could not be fulfilled. In the white region, both sugars have to be consumed to satisfy all the constraints imposed. Finally, light and dark grey regions correspond to combinations of v^u_{Glc} and v^u_{Gal} where only glucose or only galactose, respectively, is consumed, i.e. one of the two input rates is off.

Let us now analyse the trajectories followed, in the $v^u_{Glc} - v^u_{Gal}$ plane, when the sugars are consumed. In Fig. 3a, the initial point has high values of v^u_{Glc} and v^u_{Gal}. As in this region (i.e. the light grey region) only glucose is consumed, the input rate of galactose is off (v_{Gal}=0). This situation corresponds to a medium with high glucose and galactose concentrations. Glucose consumption reduces v^u_{Glc}, while v^u_{Gal} remains unchanged, the pair $\left(v^u_{Glc}, v^u_{Gal}\right)$ moving to the left, parallel to the *x*-axis. This movement continues until the region where galactose consumption is induced (dark grey region). At this point, the glucose available is low and galactose is high. Now, galactose consumption reduces v^u_{Gal}, the pair $\left(v^u_{Glc}, v^u_{Gal}\right)$ moving downwards. This shift from the glucose consuming to the galactose consuming regime is in agreement with the pattern of growth found experimentally, consisting in that the consumption of galactose is induced only after glucose is depleted from the medium [1]. On the other hand, if the initial point has low upper limit values of the two input rates, v^u_{Glc} and v^u_{Gal}, but sufficiently high values to be compatible with growth, the calculations show that both sugars are consumed (Fig. 3b, white region). This agrees with experimental studies of bacterial growth in mixtures of glucose and galactose at low concentrations where it is found that both sugars are consumed simultaneously [3, 4]. Finally, if the two sugars in the $v^u_{Glc} - v^u_{Gal}$ plane (Fig. 3b) continue to be consumed until their concentrations are very low, the amount of nutrient available is not enough to sustain growth. At this point, growth stops (black region).

As was described above, the optimisation procedure applied resulted in three patterns of sugar consumption corresponding to the three regions in the $v^u_{Glc} - v^u_{Gal}$ plane of Fig. 3 (white, light grey and dark grey regions). Inside each of these regions, there are different patterns of active fluxes. We shall analyse four flux patterns corresponding to four different combinations of values of v^u_{Glc} and v^u_{Gal}, represented by points *A*, *B*, *C* and *D* of Fig. 3a and b. Points *A* ($v^u_{Glc} = 17.0$ and $v^u_{Gal} = 15.0$) and *B* ($v^u_{Glc} = 3.0$ and $v^u_{Gal} = 15.0$) are in the region where the only carbon source used is glucose. The flux distributions in each of these

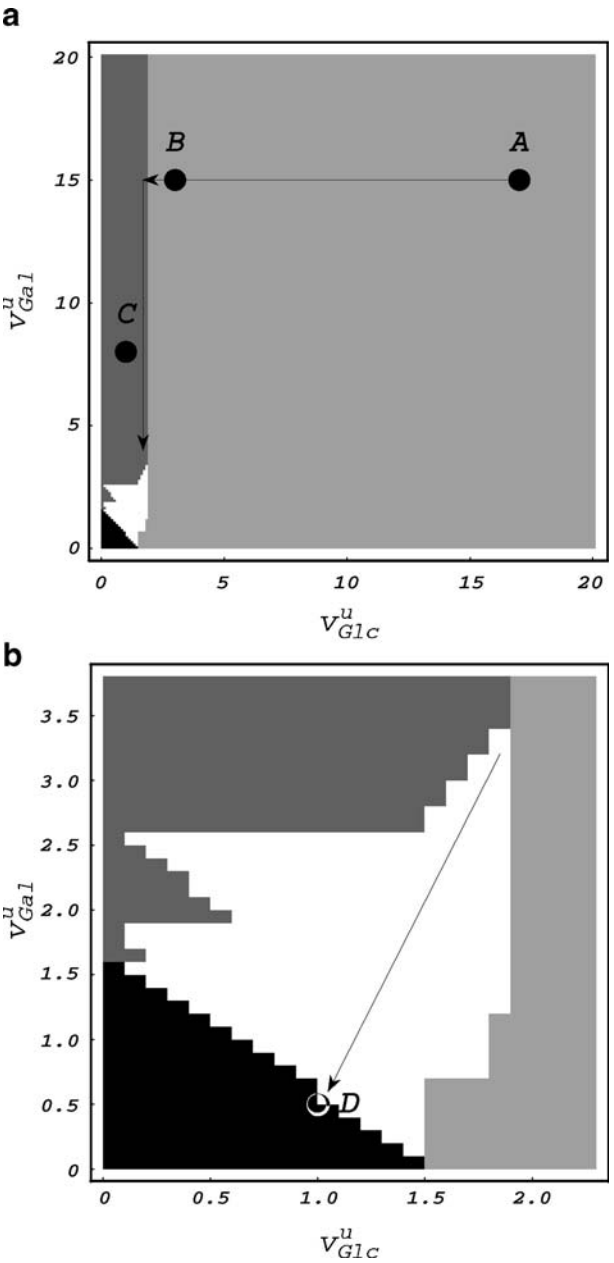

Fig. 3 Calculated patterns of sugar consumption in the $v^u_{Glc} - v^u_{Gal}$ plane. **a** Represents the whole plane and **b** an amplification of the bottom-left corner of **a**. *Black*, *white*, *light grey* and *dark grey* regions correspond to the following patterns: no sugar consumption, simultaneous consumption of glucose and galactose, glucose consumption only and galactose consumption only, respectively

two situations are shown in Fig. 4a and b, respectively. In the environmental condition of point *A*, glucose is fermented and no oxygen is consumed (although the calculations were performed in aerobic conditions, i.e. with an upper limit for the oxygen input flux of 10). When the maximal input of glucose is $v^u_{Glc} = 3.0$, the respiratory chain is turned on and oxygen is consumed. This transition from fermentation to respiration when the concentration of glucose is reduced in aerobic conditions is found experimentally. The

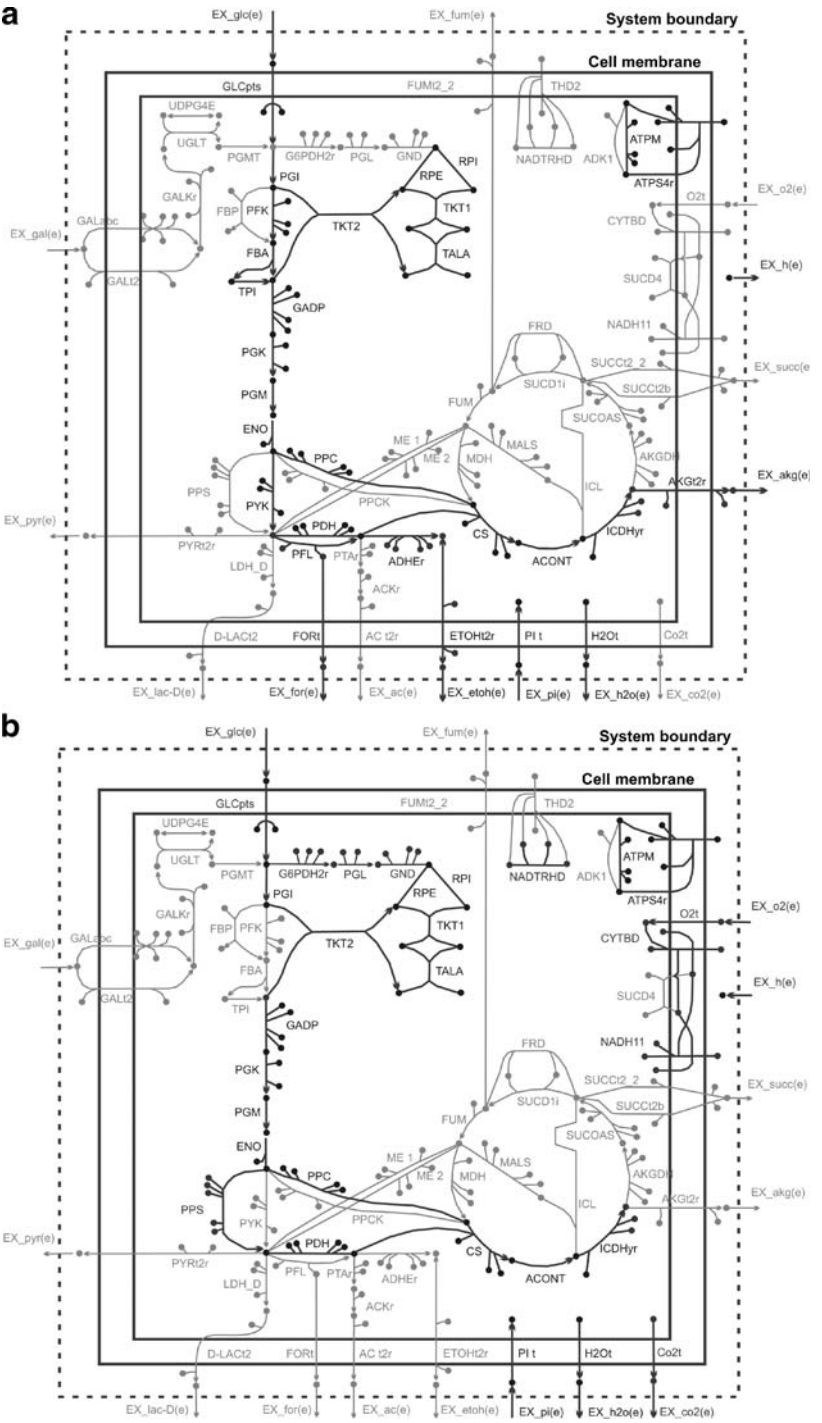

Fig. 4 Diagrammatic representation of three flux patterns of Fig. 3. Active reaction steps are in black and inactive in grey. **a** Represents point *A*, $v^{u}_{Glc} = 17.0$ and $v^{u}_{Gal} = 15.0$, having 30 steps on; **b** represents point *B*, $v^{u}_{Glc} = 3.0$ and $v^{u}_{Gal} = 15.0$, with 31 steps on and **c** represents point *D*, $v^{u}_{Glc} = 1.0$ and $v^{u}_{Gal} = 0.5$, showing 43 steps on

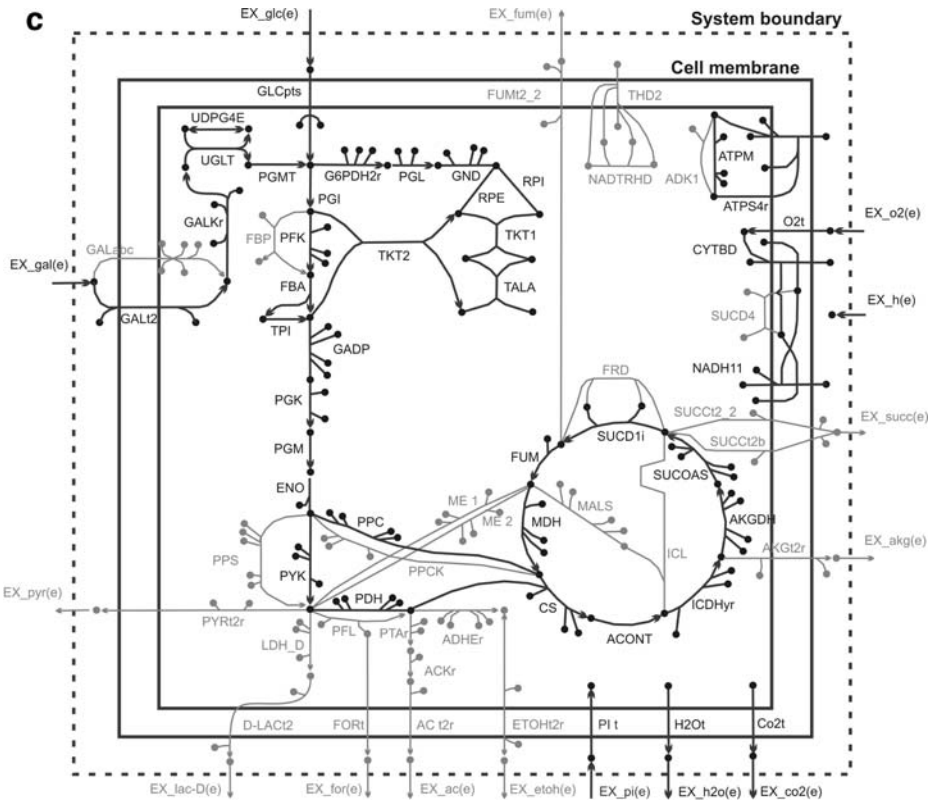

Fig. 4 (continued)

inhibition of respiration at high glucose concentrations, when the *E. coli* is grown aerobically, is called the bacterial Crabtree effect [22, 23]. In the dark grey region, the amount of glucose is too low, and the system induces galactose consumption. At point *C* ($v^u_{Glc} = 1.0$ and $v^u_{Gal} = 8.0$), belonging to this region, galactose is degraded by a respiration regime and oxygen is consumed (diagram not shown). Finally, point *D* ($v^u_{Glc} = 1.0$ and $v^u_{Gal} = 0.5$) is in the white region (very close to the black region) where consumption of both sugars is necessary to achieve the minimum growth rate. In the pattern of active fluxes corresponding to this point of the plane, the respiratory chain and all the Krebs cycle steps are on (see Fig. 4c). In Fig. 3, the transition from pure fermentation to respiro-fermentation in the light grey region occurs for v^u_{Glc} between 16.1 and 16.0 and in the dark grey region for v^u_{Gal} between 13.9 and 13.8.

The patterns of active fluxes obtained by the optimisation procedure strongly depend on the main target, namely the maximisation of v_{growth} subject to the condition that the number of active reaction steps is equal to the minimum previously calculated. For instance, the point $v^u_{Glc} = 17$ and $v^u_{Gal} = 0$ in the two-level optimisation results in $n_{\min}$=30 and v_{growth}=0.56, glucose being degraded by fermentation. However, for the same point of the $v^u_{Glc} - v^u_{Gal}$ plane, if n is minimised but v_{growth} is not maximised, there is another solution with $n_{\min}$=30 and v_{growth}=0.31, but in this situation, glucose is degraded by respiration. As a consequence, prediction of the Crabtree effect (i.e. inhibition of respiration at high glucose concentrations) requires the application of the two-level optimisation criterion. Another example where the result of optimisation depends on the application of the main target is the following. At the

point $v^{\mu}_{Glc} = 1.9$ and $v^{\mu}_{Gal} = 2.0$, the two-level algorithm gives $n_{\min}$=34 and v_{growth}=0.25, and the input fluxes of the sugars are: v_{Glc}=1.90 and v_{Gal}=1.62. In contrast, if only the secondary target is optimized, there is a solution with the same number of steps, $n_{\min}$=34, and lower v_{growth}=0.11, where only glucose is consumed (v_{Glc}=1.90 and v_{Gal}=0). In this case, the lack of maximisation of v_{growth} can even change the coarse pattern of sugar consumption (represented in Fig. 3). But the maximisation of v_{growth} alone (without previous minimisation of number of active reaction steps) results in glucose and galactose always simultaneously consumed, when both are present, which is not in agreement with experimental findings (results not shown).

5 Discussion

In this contribution, we calculated the patterns of utilisation of mixtures of glucose and galactose by *E. coli*, optimising the flux distribution constrained by the stoichiometry of its core metabolic network. The two-level target of optimisation was the minimisation of the number of active reactions (subject to attaining a growth rate higher than a lower limit) and subsequent maximisation of the growth rate (subject to the number of active reactions being equal to the minimum previously calculated). The main arguments in favour of the biological significance of this mathematical procedure are as follows. The central target of evolution is the maximisation of growth rate. In our calculations, this goal is achieved in two successive levels. In the first level, the minimisation of the number of steps, by elimination of those that do not contribute to growth, releases nutrient and protein that become available to be used to increase growth rate. In the second level, the growth rate of the network with reduced number of steps, which has been obtained in the first level, is maximised. Using this unique two-level criterion, we were able to obtain by calculation four fundamental behaviours found experimentally. First, the calculations at high glucose concentration with oxygen available predicted that this sugar is degraded by fermentation (independently of the amount of galactose present). This is in agreement with the experimental finding that respiration is inhibited at high glucose concentrations when *E. coli* is grown aerobically, which is called the bacterial Crabtree effect [22, 23]. Second, if we decrease the concentration of glucose, mimicking glucose consumption, we predicted that there is a point where sugar metabolism switches to respiration. Experimentally, when glucose is consumed aerobically, a sufficiently low concentration is reached where the system turns on respiration [23]. Third, if we continue to decrease the concentration of glucose, the model predicted that galactose replaces glucose as carbon source. In practice, when the system is depleted of glucose, galactose utilisation is genetically induced [1]. Fourth, when both sugars are present in low concentrations, the calculations predicted that they are consumed simultaneously, coinciding with what is found experimentally under these conditions [3, 4].

Regarding the predictions of the coarse patterns of response, represented in Fig. 3, there is one aspect where there appears to be a discrepancy with experimental evidence. In the calculated patterns, if glucose is consumed in the presence of a high concentration of galactose, its concentration decreases until it reaches certain value where the pathway for galactose consumption switches on and the pathway for glucose consumption switches off. In *E. coli*, the enzymes for glucose degradation have been described as constitutive, i.e. they are produced independently of the concentration of glucose that is present. According to this experimental finding, our dark grey region (Fig. 3) would have to be painted in white. It is of course also possible that at low glucose and high galactose concentrations,

there is a still unknown mechanism that reduces the amount of enzyme protein for glucose incorporation. This could operate by inhibiting enzyme synthesis or producing specific protein degradation (e.g. by ubiquitin-mediated proteolysis which was shown to be involved in the response to stress and extracellular effectors, among other responses, [24]). The existence of such a mechanism would be advantageous for the bacteria because the reduction in the concentration of the proteins for glucose degradation would leave more space for proteins degrading galactose or other processes which are relevant at low glucose concentrations.

The metabolic network with 69 internal reactions used in this work (the "core") is a condensed version of the sugar degradation metabolism of the genome-scale *E. coli* reconstruction. The number of internal reactions of the genome-scale network model has been increasing during the past years, starting with 720 reactions in 2000 and arriving at 2,077 reactions in 2007 [7, 25]. We have determined the patterns of sugar consumption for the genome-scale network of 720 reaction steps (not shown). The preliminary results obtained are very similar to those represented for the core in Fig. 3. The two patterns differ only in the shape of the boundary between the white region and the grey regions. The large number of variables in the genome-scale networks constitutes a challenge to the mixed integer linear programming algorithms available, which may halt after a prespecified computation time before converging to an optimal solution. As a consequence, we have found a few inconsistencies in the results obtained with these larger networks. For example, for some points in the $v^{u}_{Glc} - v^{u}_{Gal}$ plane, the result of increasing the values of v^{u}_{Glc} and/or v^{u}_{Gal} was to obtain solutions with a higher number of steps (which then are clearly suboptimal). Although these appear to be only minor inconsistencies, we preferred to use a network with a much smaller number of variables (i.e. the core) where the software solver can be run up to solution optimality. The complex shape obtained for the boundary of the white region, represented in Fig. 3b, corresponds to slightly different consistent solutions. In addition, the core has the advantage that it is easier to represent diagrammatically the processes that turn on and off when the external nutrients are changed. On the other hand, the core represents only a small fraction of the reactions (less than 10%) of the genome-scale network. Therefore, even if its coarse pattern of nutrient consumption is very similar to that of the full network, the detailed patterns of flux distribution may show deviations. In this respect, we found that in most of the nutrient conditions where there is fermentation, ethanol, formate, lactate and/orα-ketoglutarate are the by-products only in some small regions of the $v^{u}_{Glc} - v^{u}_{Gal}$ plane (e.g. in the point $v^{u}_{Glc} = 1.8$ and $v^{u}_{Gal} = 1.0$) with acetate being excreted. This contrasts with what is found experimentally when *E. coli* is grown in glucose, acetate being usually a product of fermentation [23]. There is strong experimental evidence in the yeast *Saccharomyces cerevisiae* that the switch from anaerobic growth to aerobic respiration upon depletion of glucose is correlated with widespread changes in gene expression. In DNA microarray studies, it was found that from 6,400 distinct DNA sequences, more than 2,000 suffered increases or decreases in their levels by factors greater that two, as glucose was progressively depleted from the growth media during the course of the experiment [26]. It is, therefore, not surprising that our simplified approach, consisting in 69 reactions only, does not reproduce the detailed patterns of flux distribution found experimentally.

As we have discussed above, the number of active reaction steps can be interpreted as the number of genes that are expressed in certain environmental conditions. Therefore, this approach assumes that the relevant secondary target is associated with what is turned on and what is turned off at the genetic level. An alternative function that could be used as the target of minimisation is the sum of the metabolic fluxes. This function is directly related to

the "effort" that the cell has to make in terms of energy and external resources to produce the enzymes and membrane transporters (see for example [13]). In essence, the relevant secondary target could represent the minimisation of the total amount of protein that the cell has to produce. We have minimised the alternative function for the core of *E. coli* (results not shown). Regarding the coarse patterns of response, represented in Fig. 3, the only difference found is that the dark grey region is now white (both sugars are consumed), which would be in agreement with the enzymes for glucose metabolism being constitutive. But analysis of the detailed patterns of flux distribution show that the only mode of functioning is respiro-fermentation. Therefore, the inhibition of respiration at high glucose concentrations in aerobic conditions (i.e. the bacterial Crabtree effect) cannot be explained using as target the minimisation of the sum of the fluxes.

The patterns of sugar consumption obtained by calculation in this work make sense if glucose is frequently an abundant component of the medium in which the cells grow. For microorganisms living in environments where glucose is not the most abundant sugar available, selective pressure would favour regulatory mechanisms where the enzymes for glucose metabolism are not constitutive. In these situations, one would expect that the bacteria prefer degrading the sugar that is normally present in the environment, the "common" sugar. This is the case for the bacteria *Streptococcus thermophilus* whose natural habitat is milk where lactose is the most abundant sugar available. When this bacteria is grown in a mixture of glucose and lactose, it first consumes lactose until its habitat is depleted of this sugar and only afterwards does it start to consume glucose [27]. The logic of this strategy is clear. There is no use in synthesising the shortest pathway (for glucose incorporation) if its substrate is not present. The best strategy is to express the pathway used to degrade the common sugar and only after this sugar has been consumed to degrade glucose, if this is present. Note that this is the strategy that minimises the number of active reaction steps because in the absence of substantial amounts of glucose (the most common situation for these bacteria), the solution with lower number of steps is to turn on only the pathway for degrading the sugar commonly available.

Acknowledgements The authors acknowledge technical support in preparing preliminary computational tools from Pablo Berger. LA is grateful for funding from Comisión Sectorial de Investigación Científica de la Universidad de la República (CSIC, Montevideo). HC and LA acknowledge the support from Programa de Desarrollo de las Ciencias Básicas (PEDECIBA, Montevideo).

Appendix 1

A. Reactions

Abbreviation	Official reaction name	Equation
ACKr	Acetate kinase	$[c]: ac + ATP \Leftrightarrow actp + ADP$
ACONT	Aconitase	$[c]: cit \Leftrightarrow icit$
ACt2r	Acetate reversible transport via proton symport	$ac[e] + h[e] \Leftrightarrow ac[c] + h[c]$
ADHEr	Acetaldehyde dehydrogenase	$[c]: accoa + (2)h + (2)\,NADH \Leftrightarrow coa + Et\,OH + (2)NAD$
ADK1	Adenylate kinase	$[c]: AMP + ATP \Leftrightarrow (2)\,ADP$

Abbreviation	Official reaction name	Equation
AKGDH	2-Oxoglutarate dehydrogenase	$[c]: akg + coa + NAD \rightarrow CO_2 + NADH + succoa$
AKGt2r	2-Oxoglutarate reversible transport via symport	$akg[e] + h[e] \Leftrightarrow akg[c] + h[c]$
ATPM	ATP maintenance requirement	$[c]: ATP + H_2O \rightarrow ADP + h + pi$
ATPS4r	ATP synthase (four protons for one ATP)	$adp[c] + (4)\, h[e] + pi[c] \Leftrightarrow ATP[c] + (3)\, h[c] + H_2O[c]$
CO2t	CO_2 transporter via diffusion	$CO_2[e] \Leftrightarrow CO_2[c]$
CS	Citrate synthase	$[c]: accoa + H_2O + oaa \rightarrow cit + coa + h$
CYTBD	Cytochrome oxidase bd (ubiquinol-8: 2 protons)	$(2)\, h[c] + (0.5)\, O_2[c] + q8h2[c] \rightarrow (2)\, h[e] + H_2O[c] + q8[c]$
D-LACt2	D-lactate transport via proton symport	$h[e] + lac - D[e] \Leftrightarrow h[c] + lac - D[c]$
ENO	Enolase	$[c]: 2pg \Leftrightarrow H_2O + pep$
ETOHt2r	Ethanol reversible transport via proton symport	$ETOH[e] + h[e] \Leftrightarrow ETOH[c] + h[c]$
EX_ac(e)	Acetate exchange	$[e]: ac \Leftrightarrow$
EX_akg(e)	2-Oxoglutarate exchange	$[e]: akg \Leftrightarrow$
EX_co2(e)	CO2 exchange	$[e]: CO_2 \Leftrightarrow$
EX_etoh(e)	Ethanol exchange	$[e]: ETOH \Leftrightarrow$
EX_for(e)	Formate exchange	$[e]: for \Leftrightarrow$
EX_fum(e)	Fumarate exchange	$[e]: fum \Leftrightarrow$
EX_glc(e)	D-Glucose exchange	$[e]: glc - D \Leftrightarrow$
EX_h(e)	H^+ exchange	$[e]: h \Leftrightarrow$
EX_h2o(e)	H_2O exchange	$[e]: H_2O \Leftrightarrow$
EX_lac-D(e)	D-lactate exchange	$[e]: lac - D \Leftrightarrow$
EX_o2(e)	O_2 exchange	$[e]: O_2 \Leftrightarrow$
EX_pi(e)	Phosphate exchange	$[e]: pi \Leftrightarrow$
EX_pyr(e)	Pyruvate exchange	$[e]: pyr \Leftrightarrow$
EX_succ(e)	Succinate exchange	$[e]: succ \Leftrightarrow$
FBA	Fructose-bisphosphate aldolase	$[c]: fdp \Leftrightarrow dhap + g3p$
FBP	Fructose-bisphosphatase	$[e]: fdp + H_2O \rightarrow f6p + pi$
FORt	Formate transport via diffusion	$for[e] \Leftrightarrow for[c]$
FRD	Fumarate reductase	$[c]: fadh2 + fum \rightarrow fad + succ$
FUM	Fumarase	$[c]: fum + H_2O \Leftrightarrow mal - L$
FUMt2_2	Fumarate transport via proton symport (2 H)	$fum[e] + (2)h[e] \rightarrow fum[c] + (2)h[c]$
G6PDH2r	Glucose 6-phosphate dehydrogenase	$[c]: g6p + NADP \Leftrightarrow 6pgl + h + NADPH$
GALKr	Galactokinase	$[c]: ATP + gal \leftarrow ADP + gal1p + h$
GALabc	D-Galactose transport via ABC system	$ATP[c] + gal[e] + H_2O[c] \rightarrow ADP[c] + gal[c] + h[c] + pi[c]$
GALt2	D-Galactose transport in via proton symport	$gal[e] + h[e] \rightarrow gal[c] + h[c]$
GAPD	Glyceraldehyde-3-phosphate dehydrogenase	$[c]: g3p + NAD + pi \Leftrightarrow 13\,dpg + h + NADH$
GLCpts	D-Glucose transport via PEP:Pyr PTS	$glc - D[e] + pep[c] \rightarrow g6p[c] + pyr[c]$

Abbreviation	Official reaction name	Equation
GND	Phosphogluconate dehydrogenase	$[c] : 6pgc + NADP \rightarrow co2 + NADPH + ru5p - D$
H2Ot	H_2O transport via diffusion	$H_2O[e] \Leftrightarrow H_2O[c]$
ICDHyr	Isocitrate dehydrogenase (NADP)	$[c] : icit + NADP \Leftrightarrow akg + CO_2 + NADPH$
ICL	Isocitrate lyase	$[c] : icit \rightarrow glx + succ$
LDH_D	D-Lactate dehydrogenase	$[c] : lac - D + NAD \Leftrightarrow h + NADH + pyr$
MALS	Malate synthase	$[c] : accoa + glx + H_2O \rightarrow coa + h + mal - L$
MDH	Malate dehydrogenase	$[c] : mal - L + NAD \Leftrightarrow h + NADH + oaa$
ME1	Malic enzyme (NAD)	$[c] : mal - L + NAD \rightarrow CO_2 + NADH + pyr$
ME2	Malic enzyme (NADP)	$[c] : mal - L + NADP \rightarrow CO_2 + NADPH + pyr$
NADH11	NADH dehydrogenase (ubiquinone-8 & 2 protons)	$(3)\, h[c] + NADH[c] + q8[c] \rightarrow (2)h[e] + NAD[c] + q8h2[c]$
NADTRHD	NAD transhydrogenase	$[c] : NAD + NADPH \rightarrow NADH + NADP$
O2t	O_2 transport (diffusion)	$O_2[e] \Leftrightarrow O_2[c]$
PDH	Pyruvate dehydrogenase	$[c] : coa + NAD + pyr \rightarrow accoa + CO_2 + NADH$
PFK	Phosphofructokinase	$[c] : ATP + f6p \rightarrow ADP + fdp + h$
PFL	Pyruvate formate lyase	$[c] : coa + pyr \rightarrow accoa + for$
PGI	Glucose-6-phosphate isomerase	$[c] : g6p \Leftrightarrow f\,6p$
PGK	Phosphoglycerate kinase	$[c] : 3pg + ATP \Leftrightarrow 13\,dpg + ADP$
PGL	6-Phosphogluconolactonase	$[c] : 6pgl + H_2O \rightarrow 6pgc + h$
PGM	Phosphoglycerate mutase	$[c] : 2pg \Leftrightarrow 3pg$
PGMT	Phosphoglucomutase	$[c] : g1p \Leftrightarrow g\,6p$
Pit	Inorganic phosphate exchange, diffusion	$pi[c] \Leftrightarrow pi[e]$
PPC	Phosphoenolpyruvate carboxylase	$[c] : CO_2 + H_2O + pep \rightarrow h + oaa + pi$
PPCK	Phosphoenolpyruvate carboxykinase	$[c] : ATP + oaa \rightarrow ADP + CO_2 + pep$
PPS	Phosphoenolpyruvate synthase	$[c] : ATP + H_2O + pyr \rightarrow AMP + (2)h + pep + pi$
PTAr	Phosphotransacetylase	$[c] : accoa + pi \Leftrightarrow actp + coa$
PYK	Pyruvate kinase	$[c] : ADP + h + pep \rightarrow ATP + pyr$
PYRt2r	Pyruvate reversible transport via proton symport	$h[e] + pyr[e] \Leftrightarrow h[c] + pyr[c]$
RPE	Ribulose 5-phosphate 3-epimerase	$[c] : ru5p - D \Leftrightarrow xu\,5p - D$
RPI	Ribose-5-phosphate isomerase	$[c] : r5p \Leftrightarrow ru\,5p - D$
SUCCt2_2	Succinate transport via proton symport (2 H)	$(2)\, h[e] + succ[e] \rightarrow (2)h[c] + succ[c]$
SUCCt2b	Succinate efflux via proton symport	$h[c] + succ[c] \rightarrow h[e] + succ[e]$
SUCD1i	Succinate dehydrogenase	$[c] : fad + succ \rightarrow fadh\,2 + fum$
SUCD4	Succinate dehyrdogenase	$[c] : fadh2 + q8 \Leftrightarrow fad + q8h2$
SUCOAS	Succinyl-CoA synthetase (ADP-forming)	$[c] : ATP + coa + succ \Leftrightarrow ADP + pi + succoa$
TALA	Transaldolase	$[c] : g3p + s7p \Leftrightarrow e\,4p + 6p$

Abbreviation	Official reaction name	Equation
THD2	NAD(P) transhydrogenase	$(2)\,h[e] + NADH[c] + NADP[c] \rightarrow (2)\,h[c] + NAD[c] + NADPH[c]$
TKT1	Transketolase	$[c] : r5p + xu5p - D \Leftrightarrow g\,3p + s7p$
TKT2	Transketolase	$[c] : e4p + xu5p - D \Leftrightarrow f\,6p + g3p$
TPI	Triose-phosphate isomerase	$[c] : dhap \Leftrightarrow g\,3p$
UDPG4E	UDPglucose 4-epimerase	$[c] : udpg \Leftrightarrow udpgal$
UGLT	UDPglucose-hexose-1-phosphate uridylyltransferase	$[c] : gal1p + udpg \Leftrightarrow g\,1p + udpgal$

Biomass production

0.2 **G6P** + 0.071 **F6P** + 0.898 **R5P** + 0.361 **E4P** + 0.129 **T3P** + 1.4996 **3PG** + 0.519 **PEP** + 2.833 **PYR** + 3.748 **AcCoA** + 1.787 **OAA** + 1.079 **alfa-KG** + 42.703 **ATP** + 18.22 **NADPH** ➔ 3.748 **CoA** + 18.22 **NADP** + 42.703 **ADP** + 42.703 **Pi** + **BIOMASS**

B. Upper and Lower Bounds of the Fluxes

1) Thirty-six reversible reactions: ACKr, ACONT, ACt2r, ADHEr, ADK1, AKGt2r, ATPS4r, CO2t, D-LACt2, ENO, ETOHt2r, FBA, FORt, FUM, G6PDH2r, GAPD, H2Ot, ICDHyr, LDH_D, MDH, PGI, PGK, PGM, PIt, PTAr, PYRt2r, RPE, RPI, SUCD4, SUCOAS, TALA, TKT1, TKT2, TPI, GALKr and PGMT are bounded by −1000 and 1000.
2) Twenty-seven irreversible reactions: AKGDH, CS, CYTBD, FBP, FRD, FUMt2_2, GND, ICL, MALS, ME1, ME2, NADH11, NADTRHD, PDH, PFK, PFL, PGL, PPC, PPCK, PPS, PYK, SUCCt2_2, SUCCt2b, SUCD1i, THD2, UDPG4E and UGLT are bounded by 0 and 1000.
3) O2t (Oxygen transport) is bounded by −1000 and 10 (reversible).
4) ATPM (ATP maintenance requirement) is bounded by 7.6 and 1000 (irreversible).
5) GROWTH (biomass production) is bounded by 0.1 and 1000 (irreversible).
6) GALabc (D-galactose transport via ABC system) and GALt2 (D-galactose transport in via proton symport): their sum is bounded by 0 and 20 (irreversible).
7) GLCpts (D-glucose transport via PEP:Pyr PTS) is bounded by 0 and 20 (irreversible).

Appendix 2: The Model

Let J be the set of all the reactions in the stoichiometry matrix $\boldsymbol{N}$ (including the 69 internal and the 15 external reactions). We then define two subsets of J: J_{int} including the 69 internal reactions and J_{ext} including the 15 external reactions.

Master optimisation problem

Objective function:

$$Z = Max \quad v_{growth}$$

Subject to:

$$\begin{aligned}
&\boldsymbol{N}\cdot\nu = o \\
&y_j\cdot\nu_j^{\mathrm{l}} \le \nu_j \le y_j\cdot\nu_j^{\mathrm{u}} \qquad \forall\, j\in J_{\mathrm{int}} \\
&\nu_j^{\mathrm{l}} \le \nu_j \le \nu_j^{\mathrm{u}} \qquad \forall\, j\in J_{\mathrm{ext}} \\
&\sum_{j\in J_{\mathrm{int}}} y_j = Z^* \\
&y_j\in\{0,1\} \\
&\nu_j\in\Re
\end{aligned}$$

where Z^* is given by the lower level optimisation problem:

$$Z^* = Min\sum_{j\in J_{int}} y_j$$

Subject to:

$$\begin{aligned}
&\boldsymbol{N}\cdot\nu = o \\
&y_j\cdot\nu_j^{\mathrm{l}} \le \nu_j \le y_j\cdot\nu_j^{\mathrm{u}} \qquad \forall\, j\in J_{\mathrm{int}} \\
&\nu_j^{\mathrm{l}} \le \nu_j \le \nu_j^{\mathrm{u}} \qquad \forall\, j\in J_{\mathrm{ext}} \\
&\nu_{\mathrm{growth}}^{\mathrm{l}} \le \nu_{\mathrm{growth}} \\
&y_j\in\{0,1\} \\
&\nu_j\in\Re
\end{aligned}$$

References

1. Monod, J.: The phenomenon of enzymatic adaptation. Growth Symposium **XI**, 223–289 (1947)
2. Egli, T.: The ecological and physiological significance of the growth of heterotrophic micro-organisms with mixtures of substrates. Adv. Microb. Ecol. **14**, 305–386 (1995)
3. Kovarova-Kovar, K., Egli, T.: Growth kinetics of suspended microbial cells: from single-substrate-controlled growth to mixed-substrate kinetics. Microbiol. Mol. Biol. Rev. **62**, 646–666 (1998)
4. Lendenmann, U., Egli, T.: Kinetic models for the growth of Escherichia coli with mixtures of sugars under carbon-limited conditions. Biotechnol Bioeng. **59**, 99–107 (1998)
5. Kacser, H., Beeby, R.: Evolution of catalytic proteins or on the origin of enzyme species by means of natural selection. J. Mol. Evol. **20**, 38–51 (1984)
6. Ortega, F., Acerenza, L.: Optimal metabolic control design. J. Theor. Biol. **191**, 439–449 (1998)
7. Edwards, J.S., Palsson, B.O.: The *Escherichia coli* MG1655 in silico metabolic genotype: its definition, characteristics, and capabilities. Proc. Natl. Acad. Sci. U. S. A. **97**, 5528–5533 (2000)
8. Palsson, B.O.: Systems biology. Properties of reconstructed networks. Cambridge University Press, New York (2006)
9. Beasley, J.E., Planes, F.J.: Recovering metabolic pathways via optimization. Bioinformatics. **23**, 92–98 (2007)
10. Burgard, A.P., Vaidyaraman, S., Maranas, C.D.: Minimal reaction sets for *Escherichia coli* metabolism under different growth requirements and uptake environments. Biotechnol. Prog. **17**, 791–797 (2001)
11. Cascante, M., Llorens, M., Meléndez-Hevia, E., Puigjaner, J., Montero, F., Marti, E.: The metabolic productivity of the cell factory. J. Theor. Biol. **182**, 317–325 (1996)
12. Henrich, R., Schuster, S., Holzhütter, H.G.: Mathematical analysis of enzymic reaction systems using optimization principles. Eur. J. Biochem. **201**, 1–21 (1991)
13. Holzhütter, H.-G.: The principle of flux minimization and its application to estimate stationary fluxes in metabolic networks. Eur. J. Biochem. **271**, 2905–2922 (2004)
14. Meléndez-Hevia, E., Waddell, T.G., Montero, F.: Optimization of metabolism: the evolution of metabolic pathways toward simplicity through the game of the pentose phosphate cycle. J. Theor. Biol. **166**, 201–220 (1994)

15. Schuetz, R., Kuepfer, L., Sauer, U.: Systematic evaluation of objective functions for predicting intracellular fluxes in *Escherichia coli*. Mol. Syst. Biol. **3**, 119 (2007)
16. Cooper, V.S., Lenski, R.E.: The population genetics of ecological specialization in evolving *Escherichia coli* populations. Nature. **407**, 736–739 (2000)
17. Acerenza, L., Graña, M.: On the origins of a crowded cytoplasm. J. Mol. Evol. **63**, 583–590 (2006)
18. Graña, M., Acerenza, L.: A model combining cell physiology and population genetics to explain the evolution of *Escherichia coli* in laboratory experiments. BMC Evol. Biol. **1**, 12 (2001)
19. Acerenza, L., Kacser, H.: Enzyme kinetics and metabolic control. A method to test and quantify the effect of enzymic properties on metabolic variables. Biochem. J. **269**, 697–707 (1990)
20. Kacser, H., Burns, J.A.: The control of flux. Symp. Soc. Exp. Biol. **27**, 65–104 (1973)
21. Levine, E., Hwa, T.: Stochastic fluctuations in metabolic pathways. Proc. Natl. Acad. Sci. U. S. A. **104**, 9224–9229 (2007)
22. Crabtree, H.G.: Observations on the carbohydrate metabolism of tumours. Biochem. J. **23**, 536–545 (1929)
23. Wolfe, A.J.: The acetate switch. Microbiol. Mol. Biol. Rev. **69**, 12–50 (2005)
24. Ciechanover, A.: The ubiquitin proteolytic system and pathogenesis of human diseases: a novel platform for mechanism-based drug targeting. Biochem. Soc. Trans. **31**, 474–481 (2003)
25. Feist, A.M., Henry, C.S., Reed, J.L., Krummenacker, M., Joyce, A.R., Karp, P.D., Broadbelt, L.J., Hatzimanikatis, V., Palsson, B.O.: A genome-scale metabolic reconstruction for *Escherichia coli* K-12 MG1655 that accounts for 1260 ORFs and thermodynamic information. Molecular Systems Biology **3**, 121 (2007)
26. DeRisi, J.L., Iyer, V.R., Brown, P.O.: Exploring the metabolic and genetic control of gene expression on a genomic scale. Science. **278**, 680–686 (1997)
27. van den Bogaard, P.T., Hols, P., Kuipers, O.P., Kleerebezem, M., de Vos, W.M.: Sugar utilisation and conservation of the gal-lac gene cluster in *Streptococcus thermophilus*. Syst. Appl. Microbiol. **27**, 10–17 (2004)