The ErTo model

A dynamical model of optimal protein allocation in *E. coli*

Abraham Lucas van Eijnatten

Research done in collaboration with and under the supervision of:

Laurens Krah

Rutger Hermsen

Major Internship report
University of Utrecht
Theoretical Biology Group

March 2022

Abstract

In order to adapt to its environment, E. coli regulates large groups of proteins in cohort using a surprisingly small number of regulators. Despite this, the regulation of protein allocation in *E. coli* often results in optimal growth rates. Previous studies have illuminated how *E. coli* distributes its resources between the various groups of resources, and what the resultant growth rates are. This has been done for wildtype *E. coli*, and for a mutant in which regulation of the group of catabolic proteins is broken. Most studies of protein allocation in *E. coli* have focused on the steady state regime. However, regulation also happens dynamically, during transitions between steady states. A recent study was able to successfully describe the dynamics of protein sectors and the growth rate under variation of the nutrient source. We derive a model of protein allocation which can potentially describe steady-state and dynamical behaviour of groups of protein and the growth rate, both in wildtype *E. coli* and the non-regulating mutant. The central assumption is that *E. coli* maximises its growth rate. Our model is able to capture steady state behaviour well, but fails in the dynamical regime. We reveal the paradoxical role of product inhibition of catabolic proteins, which within our framework is both required to predict steady state behaviour, but prevents accurate prediction of the dynamics of protein allocation. In the future our framework could shed light on the link between production inhibition and regulation in *E. coli*.

1

Layman's summary

Bacteria need to use the resources they obtain from their environments efficiently. This allows them to maximise the rate at which they grow. Being able to grow fast is an important determinant in competition between bacteria. The bacterium *E. coli* distributes its resources between various groups of functionally related proteins. In each environment *E. coli* uses regulatory proteins to modulate the expression of the various groups of proteins. Depending on the nutrient source on which *E. coli* is growing, a different fraction of the total resources needs to be allocated to each of the protein groups, to achieve an optimal growth rate. We derive a mathematical model based around differential equations which can describe how *E. coli* can regulate the distribution of resources optimally between the various groups of proteins. The model can be applied to experiments in which *E. coli* has been growing for a long time on the same nutrient source, and experiments in which a nutrient is added or removed from the growth medium. Our model is capable of describing protein allocation in both wildtype and genetically altered *E. coli*. The predictions of the model raise questions about the link between certain cellular processes and regulation. In the future we hope that our model will allow us to resolve these questions.

Contents

4

1   Introduction

In its natural environment *Escherichia coli* has to adapt its protein allocation to allow for fast growth under ever changing conditions using its regulatory network. *E. coli* does so by regulating large groups of proteins in cohort. These groups are commonly referred to as protein sectors. Proteins in a given protein sector are often controlled by the same transcriptional regulator [1, 2, 3]. The proteome can be divided into many protein sectors, each reacting differently to combinations of growth limitations [4, 5, 6]. An example of a protein sector is the C sector. Proteins belonging to the C sector are up-regulated when carbon is limiting for growth and down-regulated when nitrogen or translation are limiting for growth. The entire C sector is influenced by the transcriptional regulator CRP, which affects the expression of hundreds of genes [7]. Proteins belonging to this sector are mainly involved in catabolism or cell motility [2, 4]. Another example of a protein sector is the R sector which contains mainly proteins involved in the process of translation. These proteins are up-regulated in response to translational inhibition but down-regulated when carbon or nitrogen is limiting growth. The transcription of R sector proteins is regulated by the ppGpp molecule [8, 9]. The proteome can be divided into many more protein sectors.

Experiments have shown that *E. coli* regulates its protein allocation such that in steady-state there is an approximately linear relationship between the concentration of the protein sectors and the growth rate. Under carbon limited growth the concentration of the C sector decreases linearly with increasing growth rate (figure 1a). Under the same conditions, the concentration of the R sector follows a positive linear relationship with the growth rate (figure 1b). Such linear relationships between the steady-state concentrations of protein sectors and the growth rate are referred to as growth laws. The growth laws really represent a parametric plot between the growth rate as a function of the quality of the environment and C sector allocation as a function of the quality of the environment. That is, the growth laws encapsulate steady-state protein allocation under variation of the quality of the environment (e.g. substrate quality). Note that, despite the apparently simple form of the growth laws, it is not at all obvious that protein sectors should be linearly related to the growth rate in such a way. After all, the growth rate emerges from the combined input of a great many cellular processes and variables. The fact that regulation can nevertheless be summarised in linear relationships between the concentration of protein sectors and the growth rate is intriguing. Furthermore, it implies that models might be able to capture the regulation of protein allocation in *E. coli* using only a single internal variable.
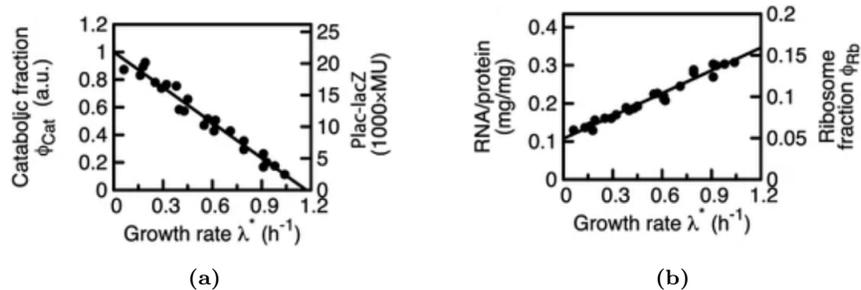
Figure 1: Regulation results in linear relationships between protein sectors and the growth rate. Figures taken from Erickson *et al* (2017). Dots are experimental measurements in different environments and the black lines are the best linear fit to the data. **a** Growth law of the C sector. LacZ activity measured in Miller units is used as a proxy for the concentration of the C sector as a whole. **b** Growth law of the R sector. RNA/protein ratio is used as a proxy for the concentration of the R sector.

The fact that E. coli regulates whole protein sectors rather than using a more specific regulatory strategy lends itself particularly well for mathematical modelling with differential equations. With relatively few variables and equations one can capture global protein allocation. In fact, some successful models of proteome allocation in E. coli coarse grain the proteome even further into only a C- R- and Q sector [3, 9, 10]. Here, the Q sector is a summary variable which models collective the behaviour of all protein sectors except the C- and R sector as being constant over time. In this simplified view, *E. coli* only has to worry about distributing its resources between the C- and R sector (figure 2). Using this coarse grained view a study done by Towbin *et al* [10] was able to successfully model the steady-state proteome allocation in both wildtype E. coli and a mutant in which the regulation of the C sector by CRP was broken.
.

(a)   **(b)**
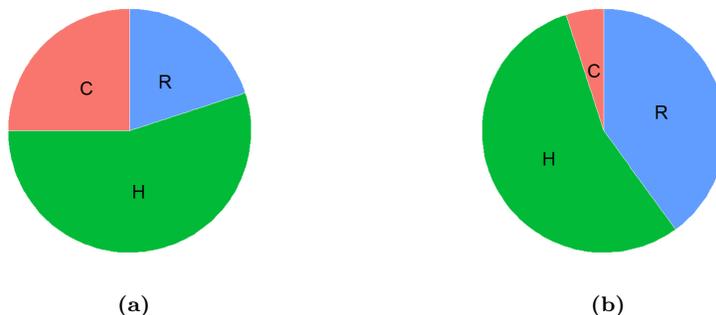
Figure 2: In a coarse grained view *E. coli* only has to distribute resources between the C- and R sector. The Q sector maintains a constant fraction of the proteome over time. **a** When carbon is very limiting (e.g. *E. coli* grows on a low quality carbon substrate), more resources are allocated towards the C sector. **b** On a high quality substrate, *E. coli* can afford to allocate more resources towards the R sector.

Most models of protein allocation in E. coli have focused on describing the steady-state protein allocation and the resultant growth rates [9, 3, 10]. However, E. coli transitions dynamically between steady-states. This out of steady-state domain can also be investigated through experiments and mathematical models. Recently, a successful study by Erickson *et al* [11] modelled the dynamics of the C- and R sectors and the growth rate during transitions between steady-states. The model was able to quantitatively match the out-of-steady-state data from these experiments. As of yet, no single model has been able to capture both the steady-state data of the mutant with broken C sector regulation from Towbin *et al* and the dynamical data from Erickson *et al* (we will explain why neither model is suitable to describe the data of the other in chapter section 3.4). Nevertheless, there appears to be no theoretical obstacle to such an endeavour, as both steady-state and dynamical behaviour likely emerge from a single regulatory network.

In this document we will attempt to derive a model describing both steady-state (wildtype and mutant) and dynamical proteome allocation in E. coli. We will do this by making a hybrid between the models of Towbin *et al* and Erickson *et al*. This will allow us to investigate to what extent the models are compatible with respect to their assumptions, derivations, interpretations and predictions. In Chapter 2 we will describe the experiments and model from Towbin *et al*. In chapter 3 we will do the same for the study by Erickson *et al*. We will describe their model in enough detail to allow the reader to reproduce their results. In chapter 4, we redo and combine the derivations of both Erickson *et al* and Towbin *et al*, seeking to apply more rigour, to derive our hybrid model. We then apply our model to the steady-state data from Towbin *et al* and the dynamical data from Erickson *et al*. We refer to our model as the ErTo model, to do justice to the two studies which inspired its formulation.

2 A steady state model can describe protein allocation in wildtype and mutant *E. coli*

In this chapter we will summarise and place into perspective the study by towbin et al. They investigated whether the growth laws are optimal (maximise the growth rate), or if they are a heuristic which ensures optimality under certain conditions but not under others. A regulatory scheme which does not always maximise the growth rate could be the result of a trade-off between the cost and precision of regulation, or between the growth rate and other fitness objectives. To investigate whether the growth laws are optimal, Towbin *et al* measured whether the C sector growth law represents the expression level of catabolic proteins that maximises the growth rate in each environment. To do this, they used a mutant in which the activity of the CRP protein was placed under experimental control. Using this mutant Towbin *et al* could vary the expression of the C sector in each environment and measure the resulting growth rates. They used a mathematical model to understand their experimental findings. From this model the authors derive conditions under which regulation is optimal.

2.1 Towbin *et al* develop a mutant that cannot regulate C sector expression

In wildtype *E. coli* C sector expression is regulated using a rather simple scheme. C sector expression is activated by the transcription factor CRP. However, CRP is only active when bound to the signalling molecule cAMP. The production of cAMP is in turn inhibited by metabolites like α-ketoacids [3]. This constitutes a negative feedback loop, where the product of catabolism inhibits the expression of catabolic genes (figure 3A). In the mutant used by Towbin *et al* both the gene responsible for producing cAMP and the gene responsible for degrading it are knocked out (figure 3B). Hence the E. coli mutant is unable to regulate the amount of cAMP in the cell. By supplying cAMP externally to the medium, the activity of CRP and hence the concentration of the C sector can be experimentally controlled.

In the mutant, the growth rate as a function of the mass fraction of the C sector on a particular medium now becomes a concave function, to which Towbin *et al* refer as an O-curve (figure 3C). By comparing the top of the O-curve with the wildtype growth rate, Towbin *et al* were able to show that for some carbon substrates the growth rate resulting from the wildtype CRP expression is indeed close to maximum. For other substrates, however, the growth laws lead to a growth rate which is significantly lower than the top of the O-curve, indicating that the growth law is sub optimal for growth on this substrate.
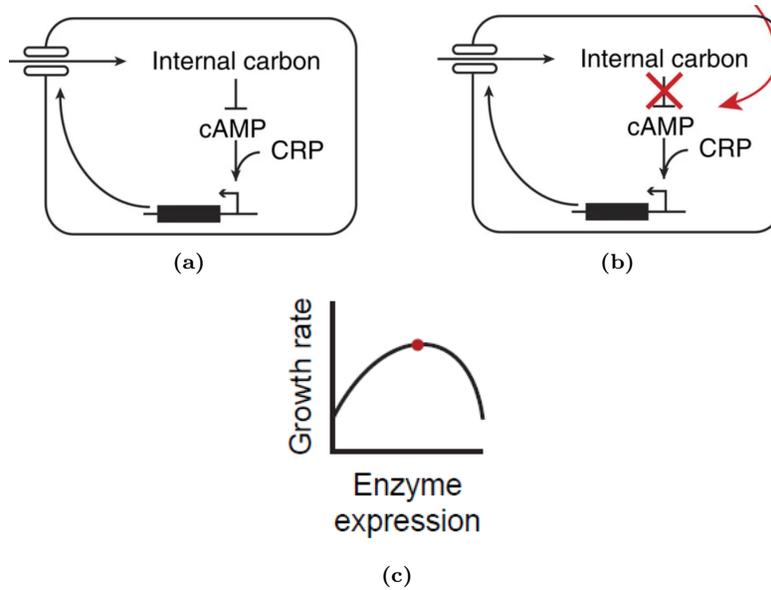
Figure 3: C sector expression is regulated through negative feedback by metabolites. Regulation is broken in the mutant. **a** CRP system in wildtype E. coli. CRP controls the expression of catabolic enzymes, which import and catabolise carbon molecules. The resulting metabolites inhibit the production of cAMP, thus inhibiting the activity of CRP. This creates a negative feedback loop. Figure taken from [10]. **b** In the mutant metabolites no longer affect the production of cAMP. Instead, cAMP is supplied externally. Figure taken from [10]. **c** Varying the external concentration of cAMP produces an O-curve, with a single growth rate optimum (red dot). Figure taken from [12].

2.2 A mathematical model can quantitatively fit to c-line and O-curve data.

2.2.1 Coarse graining the proteome into a C- R- and Q sector

Towbin *et al* sought to explain their findings using a mathematical model. They adopted the coarse grained view of the *E. coli* (depicted in figure 2), and modelled it as consisting of a C- R- and Q sector. The respective mass fractions of these sectors are denoted by ϕ_C, ϕ_R and ϕ_Q. Here, $\phi_i = \frac{M_i}{M_P}$, where M_i is the mass of the protein sector i and M_P is the total protein mass. Mass fractions are generally considered to be proportional to concentrations. This is justified because the fraction of total protein mass to total biomass is reasonably constant [11], and so is the density of the cell [13]. For this reason, the growth laws, which were originally formulated in terms of concentrations, should also hold for mass fractions, by some proportionality factor. The mass fractions sum to one, by definition:

$$\phi_C + \phi_R + \phi_Q = 1. \tag{1}$$

ϕ_C and ϕ_R are then normalised by the constant $1 - \phi_Q$ which leads to

$$\tilde{\phi}_C + \tilde{\phi}_R = 1. \tag{2}$$

Here $\tilde{\phi}_C = \frac{\phi_C}{1-\phi_Q}$ and $\tilde{\phi}_R = \frac{\phi_R}{1-\phi_Q}$.

In the Towbin *et al* model the sum of ϕ_C and ϕ_R is always constrained to be equal to $1 - \phi_Q$. We will refer to this assumption as the sum constraint. Note that this is a strong assumption which is based more on the empirical observation that models with a Q sector of constant size tend to work well than on actual biology. The Q sector is generally interpreted as a sector of housekeeping proteins, which are required in equal amounts in all environments. This view seems to be in contradiction with research indicating that the A sector (consisting of mainly anabolic genes) is regulated passively through resource competition with the C sector [14]. Despite these considerations, models with a C- R- and Q sector work have worked well in the past to describe steady-state protein allocation [3, 9].

2.2.2 A single internal variable summarises metabolism

Towbin *et al* attempted to capture metabolism from the import of carbon substrates to the production of protein in a single variable denoted x. Here, x represents a pool of metabolites central to the processes of carbon metabolism and protein production. While the precise nature of x is undetermined, possible candidates are α-ketoacids. These metabolites are precursors to amino acids. The authors assume that x inhibits the activity of C sector proteins through product inhibition, in accordance with the inhibiting effect of α-ketoacids on the CRP transcription factor. They also assume the catalytic rate of R sector proteins has a Michaelis Menten dependence on x, leading to the following equation:

$$\frac{\mathrm{d}x}{\mathrm{d}t} = \left(\rho(\kappa f(\tilde{\phi}_C) \frac{K_I}{K_I + x} - \gamma \tilde{\phi}_R \frac{x}{K_M + x} \right) = \rho \left(\kappa f(\tilde{\phi}_C) \frac{K_I}{K_I + x} - \gamma(1 - \tilde{\phi}_C) \frac{x}{K_M + x} \right). \tag{3}$$

In equation 3, κ is the maximal uptake rate on a given substrate. κ is a phenomenological parameter, which represents the rate at which a particular carbon can be converted into x. Given that in the experiments E. coli is growing under conditions on which carbon is limiting for growth, κ can be seen as representing the quality of the environment. The parameter γ represents the maximal rate of translation, K_I is the value of x at which the flux through the carbon sector is half maximal, K_M is the value of x at which the flux through the R sector is half maximal and ρ is a factor which converts from units of x to units of protein mass. The expression of the import protein is a function $f(\tilde{\phi}_C)$ of the mass fraction of the C sector. For most proteins the expression of the

import protein is proportional to the C sector such that $f(\tilde{\phi}_C) = \tilde{\phi}_C$. Towbin *et al* also consider other forms of $f(\tilde{\phi}_C)$ such as a linear relationship with a nonzero intercept $f(\tilde{\phi}_C) = \tilde{\phi}_{C,0} + \tilde{\phi}_C$ and a non monotonic function of the form $f(\tilde{\phi}_C) = \frac{\tilde{\phi}_C}{1 + (\frac{\tilde{\phi}_C}{\tilde{\phi}_{C,max}})^2}$. Here $\tilde{\phi}_{C,max}$ is the value of $\tilde{\phi}_C$ at which the value of $f(\tilde{\phi}_C)$ is highest.

The authors scale x by a factor of $\frac{1}{K_M}$ to eliminate a parameter from the model, such that

$$\frac{\mathrm{d}\hat{x}}{\mathrm{d}t} = \rho(\hat{\kappa}f(\tilde{\phi}_C)\frac{\hat{K}_I}{\hat{K}_I + \hat{x}} - \hat{\gamma}\tilde{\phi}_R\frac{\hat{x}}{1 + \hat{x}}) = \rho(\hat{\kappa}f(\tilde{\phi}_C)\frac{\hat{K}_I}{\hat{K}_I + \hat{x}} - \hat{\gamma}(1 - \tilde{\phi}_C)\frac{\hat{x}}{1 + \hat{x}}).$$
(4)

Variables carrying a circumflex have been scaled with a factor of $\frac{1}{K_M}$.

2.2.3   Defining the growth rate as the protein flux per unit of protein mass

A metabolic flux represents a rate of metabolic turnover through e.g. a protein or pathway. While Towbin *et al* do not explicitly define the fluxes through the protein sectors in their paper, they can be derived from equation 4. We will forego the derivation here but instead provide a rigorous derivation of equation 4 as a balance between the flux through the C- and the R sector in chapter 4. For now we will just define the flux through the C sector within the Towbin *et al* framework as follows:

$$J_C = \hat{\kappa}f(M_P\tilde{\phi}_C)\frac{\hat{K}_I}{\hat{K}_I + \hat{x}}.$$
(5)

Similarly J_R, the flux through the R sector (ribosomal flux) in terms of $\hat{x}$ is described by the following equation:

$$J_R = \hat{\gamma}M_P\tilde{\phi}_R\frac{\hat{x}}{1 + \hat{x}}.$$
(6)

Equation 4 describes a balance between the carbon flux which contributes to the pool of $\hat{x}$ and the ribosomal flux which drains the pool of $\hat{x}$. Both processes depend on $\hat{x}$ through product inhibition and Michaelis Menten kinetics respectively. In steady-state $\frac{\mathrm{d}\hat{x}}{\mathrm{d}t} = 0$ and there is flux balance.

Observe that equations 43 and 44 describe represent the net result of rather complicated processes. The carbon flux is determined by the flux through a range of catabolic pathways. These are each in turn comprised of many proteins which each carry their own flux at any given time. The process of catabolism is nevertheless captured in a simple equation by capturing the rate on which a

given carbon can be converted into the end product of catabolism ($\hat{x}$) with a phenomenological parameter. Beyond that the carbon flux than depends only on the size of the C sector and the effect of product inhibition by $\hat{x}$. The process of translation is described by assuming a simple Michaelis Menten relationship with $\hat{x}$. The simplicity of equations 5 and 6 stands in stark contrast with constraint based approaches like flux balance analysis, which incorporate huge amounts of data to make very complicated models of metabolism. The success of models like that of Towbin *et al* shows that the metabolism of E. coli can instead be understood on a coarse scale as emerging from only a few global assumptions.

Towbin *et al* define the growth rate λ as being equal to

$$\lambda = \gamma \tilde{\phi}_R \frac{\hat{x}}{1 + \hat{x}}. \tag{7}$$

This quantity represents the ribosomal flux per unit of protein mass. That is, Towbin *et al* define the growth rate as equal to the protein flux per unit of protein mass. This is commonly done, and is justified by noting that the density of the cell is relatively constant, and that the protein mass maintains an approximately constant fraction of the biomass. Therefore, growth in protein mass must be approximately proportional to volume growth.

2.2.4 Heuristic regulation functions determine protein flux distribution between the C- and R sector.

A central component of protein allocation models based around differential equations are the regulation functions. These functions determine what fraction of the protein flux is allocated to a particular sector. Towbin *et al* posit the functional form of their regulation functions *a priori*, rather than deriving them. They then fit these flexible functions to the growth law (The fitting procedure is described in section 2.2.7). The functional form of the regulation function of the C sector chosen by Towbin *et al* is

$$\mathcal{X}_{\tilde{C}}(\hat{x}) = \frac{\hat{K}_F}{\hat{K}_F + \hat{x}}. \tag{8}$$

Here $\mathcal{X}_{\tilde{C}}$ is the regulation function and K_F is the value of $\hat{x}$ at which resources are distributed equally between the C- and R sectors. Equation 8 behaves intuitively. In the absence of $\hat{x}$, all resources are drawn away from the R sector and allocated to the C sector. This makes sense, considering that no protein production can take place without $\hat{x}$ anyway. For very large values of $\hat{x}$, practically the entire protein flux is allocated to the production of R sector proteins. The equation for the regulation function of the R sector is then:

$$\mathcal{X}_{\tilde{R}}(\hat{x}) = 1 - \mathcal{X}_{\tilde{C}}(\hat{x}) = \frac{\hat{x}}{\hat{K}_F + \hat{x}}. \tag{9}$$

Here we used the fact that in steady-state $\mathcal{X}_i = \phi_i$. This means we can apply the sum constraint to the regulation functions.

Notice that the regulation functions described in equations 8 and 9 are a function of $\hat{x}$. This implies that the cell regulates by measuring the value of $\hat{x}$. Or to put this another way, the Towbin *et al* model assumes that E. coli infers the quality of its environment by measuring the concentration of the central metabolite, and adjusts its protein allocation accordingly. $\hat{x}$ is in this sense a proxy for $\hat{\kappa}$, which really represents the quality of the environment. Observe that here regulation of protein allocation is modelled using only a single variable. The simple steady-state relationships of the protein sectors with the growth rate (growth laws) imply that it might be possible to describe regulation using a single variable, which is central to the process of growth. Towbin *et al* do this using the central metabolite $\hat{x}$.

Figure 4 shows the regulation functions with the best fit value of K_F, and their sum. Due to the sum constraint the latter quantity is constant.
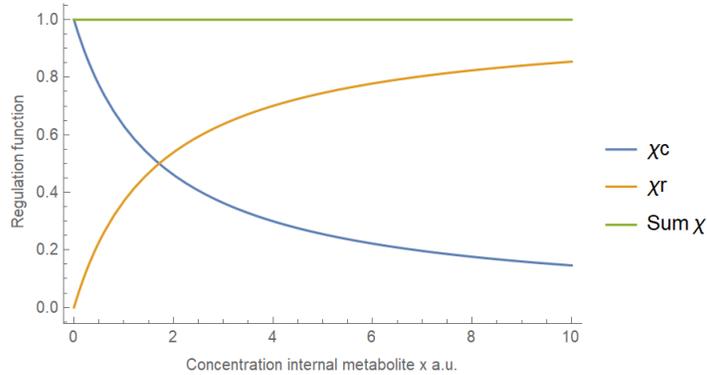
Figure 4: The heuristic regulation functions postulated by Towbin *et al* model the protein allocation between the C- and R sector.

2.2.5 Protein sector dynamics as a balance between protein production and dilution

Protein sector dynamics are modelled using equations

$$\frac{\mathrm{d}\tilde{\phi}_C}{\mathrm{d}t} = \lambda(\mathcal{X}_{\tilde{C}} - \tilde{\phi}_C) \tag{10}$$

and

$$\frac{\mathrm{d}\tilde{\phi}_R}{\mathrm{d}t} = \lambda(\mathcal{X}_{\tilde{R}} - \tilde{\phi}_R). \tag{11}$$

Here the first term $\lambda \mathcal{X}_{\tilde{i}}$ represents protein production and the second term $\lambda \tilde{\phi}_i$ dilution.

2.2.6 Deriving O-curves and a non-linear c-line from the model

The authors derive the O-curves from the model by ignoring equation 8 and treating $\tilde{\phi}_C$ as a constant. This is in accordance with the mutant in which the regulation of CRP is broken and the C sector expression is instead determined by an experimentally applied external concentration of cAMP. By substituting the constant value of $\tilde{\phi}_C$ into equation 4, equating the resulting expression to 0 and solving for $\hat{x}$ the authors calculate the steady-state value of $\hat{x}$, denoted $\hat{\bar{x}}$. By substituting $\hat{\bar{x}}$ into equation 7, Towbin *et al* derive an expression for the growth rate of the mutant as a function of the externally set C sector expression on a particular substrate:

$$\lambda(\tilde{\phi}_C) = \frac{\hat{K}_I(\tilde{\phi}_C(\kappa - \gamma) + \gamma) - \sqrt{\hat{K}_I(4(1 - \tilde{\phi}_C)\tilde{\phi}_C \kappa \gamma + \hat{K}_I(\gamma - \tilde{\phi}_C(\kappa + \gamma))^2)}}{2(\hat{K}_I - 1)}.$$

(12)

In the above equation $f(\tilde{\phi}_C) = \tilde{\phi}_C$. Towbin also derives analytical expressions for the O-curves for the other forms of $f(\tilde{\phi}_C)$ mentioned earlier; we will, however, omit these here.

Towbin *et al* derive the C-line from their model by noting that in steady-state

$$\tilde{\bar{\phi}}_C = \mathcal{X}_{\tilde{C}} = \frac{\hat{K}_F}{\hat{K}_F + \hat{\bar{x}}}.$$

(13)

The bar on top of variables signifies they are in steady-state. Solving this equation for $\hat{\bar{x}}$ leads to the following equation:

$$\hat{\bar{x}} = \hat{K}_F(\frac{1}{\tilde{\bar{\phi}}_C} - 1).$$

(14)

Substituting equation 14 into equation 7 leads to an equation for the steady-state growth rate as a function of the steady-state $\tilde{\phi}_C$:

$$\bar{\lambda} = \frac{\hat{\gamma}(1 - \tilde{\bar{\phi}}_C)^2 \hat{K}_F}{\tilde{\bar{\phi}}_C + \hat{K}_F(1 - \tilde{\bar{\phi}}_C)}.$$

(15)

Note that equation 15 is not a linear function of ϕ_C. Instead, the c-line derived by Towbin *et al* slopes up for low values of the growth rate (figure 5). Despite this, equation 15 fits the data well and is not in contradiction with the empirically observed linearity of the c-line, being approximately linear for the range of growth rates spanned by the data.
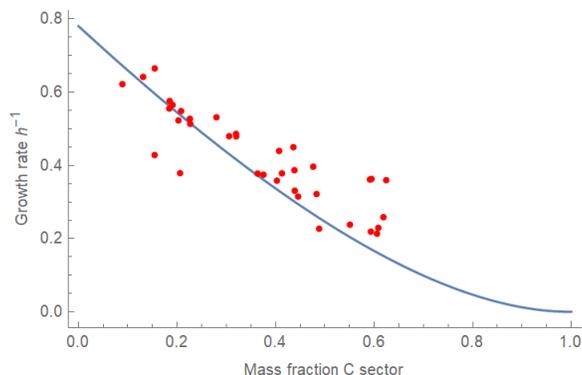
Figure 5: Towbin *et al* derive a non-linear c-line. Red dots are measurements of the steady-state relationship between λ and $\tilde{\phi}_C$ in wildtype *E. coli*. Blue line is the best fit of the c-line derived by Towbin *et al* (equation 15) to the data.

2.2.7 Towbin *et al* fit two parameters separately for each substrate

Towbin *et al* then fit the O-curves (equation 12) they have derived from their model to the experimental data they obtained by growing the mutant under varying external cAMP concentrations on several substrates. They fit $\hat{\gamma}$ and θ (which converts from measured CRP activity to $\tilde{\phi}_C$) as global parameters, and $\hat{K}_I$ and $\hat{\kappa}$ as substrate specific parameters. Whether the product inhibition constant $\hat{K}_I$ is in fact substrate specific, or has a similar effect regardless of what protein is importing the substrate is not known. Either way, Towbin *et al* allow themselves a lot of freedom by fitting two parameters per environment.

Towbin *et al* fit the c-line (equation 15) to the experimental data they obtained by letting wildtype E. coli grow in steady-state on various substrates. Because they use the value of $\hat{\gamma}$ obtained in the O-curve fit, the only remaining fitting parameter is $\hat{K}_F$.

2.3 Towbin *et al* derive and experimentally validate optimality conditions from the model

Towbin et al have shown experimentally that the growth laws do not always maximise the growth rate. We can now understand the growth laws as a heuristic which is sometimes but not always optimal. Towbin *et al* went a step further and used their model to investigate under what conditions regulation might be optimal. They showed that for their chosen regulation function regulation is optimal if $f(\tilde{\phi}_C) = \tilde{\phi}_C$ and $K_I = K_M = K_F$. Note that these optimality conditions depend on the chosen form of their regulation functions.

The authors proceeded to verify the optimality conditions derived from their model. They show that for most optimally regulated substrates, it is indeed true

that $K_I = K_M = K_F$. Towbin *et al* also show experimentally that wildtype regulation during growth on carbon substrates where $f(\tilde{\phi}_C)$ takes the form $\tilde{\phi}_{C,0} + \tilde{\phi}_C$, such as pyruvate and glycerol, is indeed sub optimal. In another experiment, the authors set a non-zero baseline expression of the sorbitol import enzyme, which in the wildtype is optimally regulated. This caused a growth rate significantly lower than the O-curve maximum, breaking optimality. The galactose import protein expression in wildtype E. coli has a non monotonic form of $f(\tilde{\phi}_C)$, leading to sub optimal growth rate. By experimentally setting $f(\tilde{\phi}_C) = \tilde{\phi}_C$, the authors were actually able to restore optimality in the wildtype. In this case E. coli appears to intentionally regulate sub optimally. Possibly this merely means that what is optimal in the lab is not optimal in a more complex environment, in which there might be competition or rapidly changing conditions.

The Towbin *et al* model is quite powerful because it is able to quantitatively match to experimental data and to make qualitative predictions regarding the principles of optimality in E. coli. It is able to do this despite being a simple model with relatively few variables and parameters, and only a single internal variable.

3 A dynamical model quantitatively describes transitions between steady states

3.1 Erickson *et al* performed out-of-steady-state experiments

A paper published in the same year by Erickson *et al* [11] investigated the growth dynamics of E. coli during transitions between steady-states. This is an entirely different domain than that analysed by Towbin *et al*, who only applied their model to steady-state data. To get out-of-steady-state data, Erickson *et al* performed upshift experiments (figure 6) in which after a period of steady-state growth on a single carbon substrate a co-utilized second carbon substrate was added. This caused *E. coli* to transition to a new steady-state with a higher growth rate. Erickson *et al* also performed downshift experiments (figure 7), in which after steady-state growth on two co-utilized carbons one of these substrates is depleted, such that *E. coli* transitioned to a new steady-state with a lower growth rate. In both upshift and downshift experiments E. coli is grown until it achieves a balanced steady-state growth before the start of the experiment. During the experiments they measured the optical density (as a proxy for the total protein mass), RNA abundance (as a proxy for relative R sector expression), LacZ expression (as a proxy for the relative C sector expression: LacZ is a catabolic protein whose expression is known to be proportional to the C sector). Erickson *et al* performed many upshift- and downshift experiments with different substrate pairs.
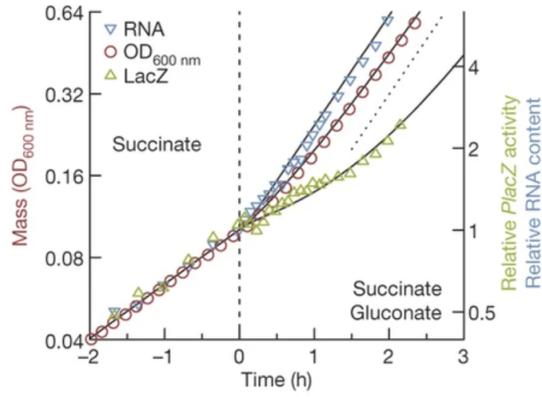
(a)
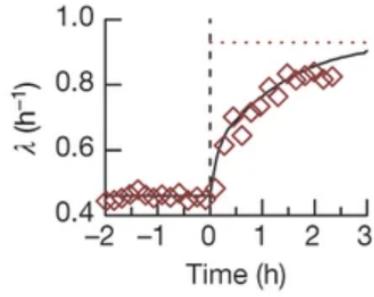
(b)
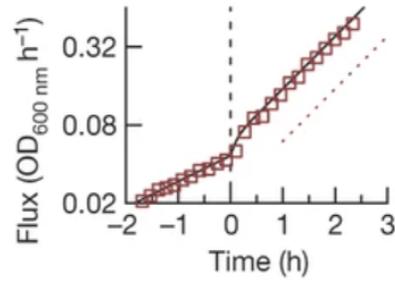
(c)

Figure 6: In upshift experiments *E. coli* transitions from steady state growth on one carbon to growth on two carbons. Figures from [11]. Dots are experimental measurements and black lines are model predictions. Growth is on succinate only until at t=0 gluconate is added to the medium. **a** Mass of total protein (red circles), R sector (blue inverted triangles) and C sector (green triangles), normalised to their respective values at $t = 0$. **b** Growth rate. **c** Biomass flux.

(a)

(b)

(c)
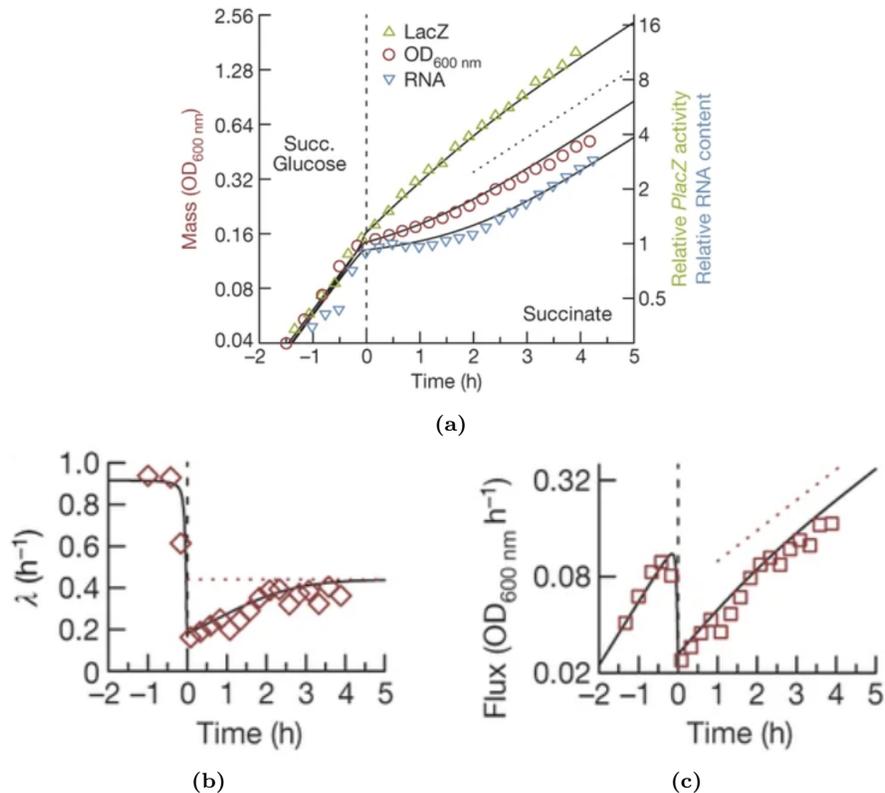
Figure 7: In downshifts *E. coli* grows in steady state on two carbons until one is depleted. Figures from [11]. Dots are experimental measurements and black lines are model predictions. Growth is on succinate and glucose until around $t = 0$ glucose is depleted and only succinate remains in the medium. **a** Mass of total protein (red circles), R sector (blue inverted triangles) and C sector (green triangles), normalised to the respective values at $t = 0$. **b** Growth rate. **c** Biomass flux.

3.2 Deriving a dynamical model of protein allocation from the growth laws

In this section we will describe the model by Erickson *et al*. We will go beyond merely positing the central equations required for interpreting their model. Instead, we will provide enough detail to perform the upshift- and downshift simulations and reproduce the results Erickson *et al* show in their paper. We have ourselves spent considerable efforts in reproducing their results and will show some of the reproductions in section 3.3.

3.2.1 Erickson *et al* do not model product inhibition or a Michaelis Menten dependence of the catabolic rate of ribosomes

Erickson *et al* model the carbon flux J_C during growth according to the following equation:

$$J_C = k_1 M_{C,1} + k_2 M_{C,2}. \tag{16}$$

Here $M_{C,1}$ and $M_{C,2}$ are the masses of two transport proteins (transporting different substrates). We assume that these transport proteins are limiting for the carbon flux. During both the upshift- and downshift experiments one substrate is present throughout the entire experiment, and one is added/depleted respectively. We refer to the former as substrate 1 and the latter as substrate 2. The flux in carbon atoms through the protein transporting substrate 1 is then $k_1 M_{C,1}$, while the flux through substrate 2 is $k_2 M_{C,2}$.

Note that equation 16 shows that Erickson *et al* do not model product inhibition, as opposed to Towbin *et al*. The carbon flux does not depend on some metabolic intermediate in this model. The authors model the uptake parameter k_i as being a function of the concentration of the substrate n_i:

$$k_i = k_{i,max} \frac{n_i}{K_{Ms,i} + n_i}. \tag{17}$$

Here $k_{i,max}$ is the maximal uptake rate of the protein responsible for transporting substrate i and $K_{Ms,i}$ is the value of n_i at which the uptake rate of the transport protein is half maximal. Because substrate 1 is present in saturating amounts during both upshifts and downshifts, equation 17 reduces to $k_1 = k_{1,max}$. In upshifts, the substrate which is added at the time of the shift, denoted by $t = 0$, is also present in saturating conditions, such that equation 17 becomes $k_2 = k_{2,max}$. However, during downshifts the substrate which is being depleted is no longer present in saturating amounts shortly before $t = 0$. We therefore need to simulate n_2 in order to calculate the value of k_2 during downshifts.

The ribosomal flux, which can be interpreted as the rate of protein production, is modelled as being always in flux balance with the carbon flux:

$$J_R = \frac{\mathrm{d}M_P}{\mathrm{d}t} = \alpha J_C = \alpha(\kappa_1 M_{C,1} + \kappa_2 M_{C,2}). \tag{18}$$

Here M_P is the total amount of protein and α is a factor which converts from number of carbon atoms to units of protein mass. Note that equation 18 shows that Erickson *et al* assume no ribosome saturation occurs. Equation 18 means that the carbon- and ribosomal flux, when written in the same units, are always equal. This is called a flux balance assumption. Erickson *et al* mention in their paper that equation 18 can be understood through a central metabolite whose concentration adapts on much faster timescales than the protein masses. That is, Erickson *et al* implicitly make a quasi-steady-state assumption for the

metabolite x.

Erickson *et al* show experimentally (See extended data figure 1A and B from the paper) that the biomass flux J_B is approximately proportional to the protein flux such that

$$J_B = \frac{\mathrm{d}M_B}{\mathrm{d}t} = \frac{J_R}{\tau}. \tag{19}$$

Here M_B is the total biomass and $\tau \approx 0.7$ is a factor which converts from biomass to protein mass.

Erickson defines the growth rate as

$$\lambda = \frac{1}{M_B}\frac{\mathrm{d}M_B}{\mathrm{d}t} = \frac{1}{M_B}J_B. \tag{20}$$

This definition of the growth rate is equivalent to the definition from Towbin *et al* as $\frac{J_B}{M_B} = \frac{\frac{J_R}{\tau}}{\frac{M_R}{\tau}} = \frac{J_R}{M_R}$. When simulating the system, the biomass can be calculated by integrating equation 19. The growth rate then follows from equation 20.

3.2.2 The regulation functions are derived from the growth laws

Whereas Towbin *et al* posited a functional form of the regulation functions of the C- and R sector *a priori*, Erickson *et al* derive the regulation functions from the growth laws. The growth law for the mass fraction $\phi_{C,i}$ of the catabolic protein responsible for the transport of substrate i is given by

$$\bar{\phi}_{C,i}(\bar{\lambda}) = \phi_{C,i}^M(1 - \frac{\bar{\lambda}}{\lambda_C}). \tag{21}$$

Here $\phi_{C,i}^M$ represents the maximal proportion of the C sector, scaled between 0 and 1, that is dedicated to the specific transport protein. λ_C is the maximum steady-state growth rate. In equation 21 it is assumed that the expression of the transport protein is proportional to the C sector.

The growth law of the R sector is given by

$$\bar{\phi}_R(\bar{\lambda}) = \phi_{R,0} + \frac{\bar{\lambda}}{\gamma}. \tag{22}$$

Here $\phi_{R,0}$ is the value of ϕ_R in an environment which does not sustain growth.

While the regulation functions from Towbin *et al* are a function of x, Erickson *et al* derive a regulation function which depends on the translational activity σ, where σ is defined as

$$\sigma = \frac{J_R}{M_R}. \tag{23}$$

Here M_R is the mass of the R sector (which consists of ribosomal proteins). Thus σ can be interpreted as the average translation rate per ribosome. Because σ depends on the concentration of precursors in the cell, σ is implicitly a function of x. However, the exact relation between σ as defined in Erickson *et al* and x as defined in Towbin *et al* is unclear, because Erickson *et al* do not explicitly model a Michaelis Menten dependence of the catalytic rate of ribosomes.

Since in steady-state $J_R = \bar{\lambda} M_P$, the following equation holds:

$$\bar{\sigma} = \frac{\bar{\lambda} M_P}{\phi_R M_P} = \frac{\bar{\lambda}}{\phi_R} = \frac{\bar{\lambda}}{\phi_{R,0} + \frac{\bar{\lambda}}{\gamma}}. \tag{24}$$

This equation can be inverted to get $\bar{\lambda}$ as a function of $\bar{\sigma}$:

$$\bar{\lambda}(\bar{\sigma}) = \frac{\bar{\sigma} \phi_{R,0}}{1 - \frac{\gamma}{\bar{\sigma}}}. \tag{25}$$

The authors assume that the cell measures σ and regulates as if the cell is in steady-state. This way of regulating has been suggested to be optimal in a recent theoretical paper [15]. It might also simply be hard for the cell to infer if it is in steady-state, such that simply regulating such that it at least ends up in the proper steady-state is the best the cell can do. A more specific way of phrasing the approach taken by Erickson *et al* is that even out-of-steady-state the cell allocates the ribosomal flux such that in steady-state it reproduces the growth laws. Mathematically this means that the authors derive the regulation functions of the C- and R sector by substituting equation 25 into the respective growth law, and then substituting σ for $\bar{\sigma}$ in the resulting equations. We will denote the substitution of variable x by variable y with $[y/x]$. The regulation functions become

$$\mathcal{X}_R(\sigma) = \bar{\phi}_R(\bar{\lambda}(\bar{\sigma}))[\sigma/\bar{\sigma}] = \frac{\phi_{R,0}}{1 - \frac{\sigma}{\gamma}} \tag{26}$$

and

$$\mathcal{X}_C(\sigma) = \frac{\bar{\phi}_C(\bar{\lambda}(\bar{\sigma}))[\sigma/\bar{\sigma}]}{\phi_{C,i}^M} = 1 - \frac{\sigma}{\lambda_C} \frac{\phi_{R,0}}{1 - \frac{\sigma}{\gamma}} = 1 - \frac{\sigma}{\lambda_C} \mathcal{X}_R(\sigma) \tag{27}$$

(figure 8). We can calculate the maximal translational activity σ_C by solving $\bar{\lambda}(\bar{\sigma}) = \lambda_C$ for σ. This results in

$$\sigma_C = \frac{1}{\frac{1}{\gamma} + \frac{\phi_{R,0}}{\lambda_C}}. \tag{28}$$

In steady-state $\mathcal{X}_R$ ranges between $\phi_{R,0}$ when $\sigma = 0$, and approximately 0.16 when $\sigma = \sigma_C$. $\mathcal{X}_C$ ranges in steady-state between 0 when $\sigma = \sigma_C$ and 1 when $\sigma = 0$. During growth transitions σ could become higher than σ_C such that $\mathcal{X}_C$ becomes negative, but this situation did not occur for any of the experiments

performed and modelled by Erickson *et al.* Figure 8 shows the regulation functions derived by Erickson *et al*, and their sum.
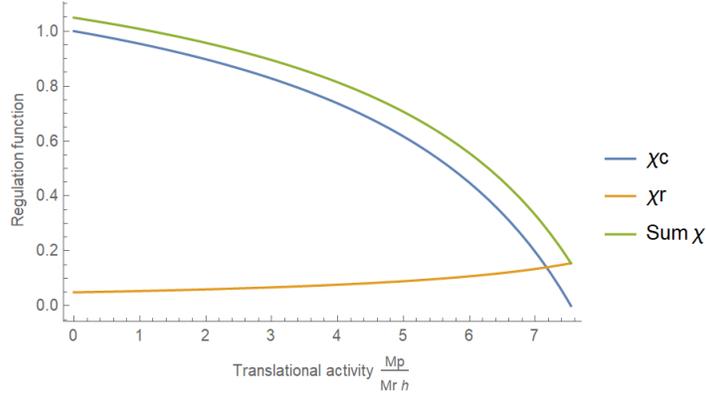

Figure 8: The regulation functions from the Erickson *et al* model plotted from 0 to σ_C.

Note that $\mathcal{X}_C$ as derived by Erickson *et al* and depicted in figure 8 cannot be interpreted as the actual regulation function of the C sector. Under severe carbon limitation such that $\bar{\lambda} = 0$ (and therefore $\sigma = 0$), $\mathcal{X}_C = 1$. In reality, however, the existence of the Q sector means not all of the protein flux is allocated to the C sector under such conditions. This discrepancy is the result of the fact that $\mathcal{X}_C$ is derived from a growth law which is scaled between 0 and 1 (i.e. $1 - \frac{\bar{\lambda}}{\lambda_C}$) whereas the real c-line must have a y-axis intercept with $\phi_C < 1$. However, since both the real c line and $1 - \frac{\bar{\lambda}}{\lambda_C}$ are linear with the same x-axis intercept, we can at least say that $\mathcal{X}_C$ is derived from a growth law which is proportional to that of the C sector. From equation 27 it then follows that $\mathcal{X}_C$ is proportional to the actual regulation function of the C sector. These complications mean that it is not possible to retrieve the mass fraction of the C sector during simulations of the model by Erickson *et al* without additional information on the actual y-axis intercept of the growth law of the C sector.

While the above makes the sum of the regulation functions (figure 8) hard to interpret, it is from the derivation clear that $\mathcal{X}_C + \mathcal{X}_R$ is not explicitly constrained to be constant. Without the sum constraint assumption (equation 1) Erickson *et al* allow themselves more freedom to describe the dynamics of the system.

3.2.3 The uptake rates can be calculated using a phenomenological lumped parameter

The authors describe the time evolution of the protein masses using differential equations. These are of the form

23

$$\frac{\mathrm{d}M_{C,i}}{\mathrm{d}t} = h_i \mathcal{X}_C J_R = h_i \mathcal{X}_C \alpha (k_1 M_{C,1} + k_2 M_{C,2}) \tag{29}$$

and

$$\frac{\mathrm{d}M_R}{\mathrm{d}t} = \mathcal{X}_R J_R. \tag{30}$$

Here $h_i = \begin{cases} 0 & \text{if substrate i is absent} \\ \phi_{C,i}^M & \text{otherwise} \end{cases}$ and M_R is the mass of the R sector. This means that the transport protein $M_{C,i}$ is not produced when its cognate carbon substrate is absent, whereas the part of the protein flux allocated to it becomes proportional to the flux allocated towards the C sector when its substrate is present. Note that this implies an additional layer of regulation for the transport proteins, in addition to the regulation by the CRP transcription factor to which all C sector proteins are subject. Erickson *et al* assume that most proteins of the C sector fluctuate up and down proportional to the C sector at all times. Meanwhile, the transport proteins are subject to additional regulation but constitute such a small proportion of the C sector that they do not significantly affect the mass of the C sector as a whole .

Observe that Erickson *et al* simulate the masses of the protein sectors, rather than their mass fraction. However, by writing $\frac{\mathrm{d}\frac{M_i}{M_P}}{\mathrm{d}t}$ in terms of $\frac{\mathrm{d}M_P}{\mathrm{d}t}$ and $\frac{\mathrm{d}M_i}{\mathrm{d}t}$ using the quotient rule one can easily rewrite equations 29 and 30 in terms of mass fractions to end up with equations 10 and 11 from Towbin *et al* (except for the initial h_i factor). Therefore, the difference between the two approaches is mostly one of convention.

The values of the parameters $\phi_{C,i}^M$, k_i and α are not specified individually by Erickson *et al*. However, they use the growth laws to derive an expression for the lumped parameter $\kappa_i = \phi_{C,i}^M \alpha k_i$. They do this by noting that in steady-state $\frac{\mathrm{d}M_{C,i}}{\mathrm{d}t} = \bar{\lambda} M_{C,i}$ and $\phi_{C,i}^M \mathcal{X}_C = \bar{\phi}_{C,i}(\bar{\lambda})$. Therefore, equation 29 in steady-state on a single substrate i becomes:

$$\bar{\lambda} M_{C,i} = \bar{\phi}_{C,i}(\bar{\lambda}) \alpha k_1 M_{C,i} = \phi_{C,i}^M (1 - \frac{\bar{\lambda}}{\lambda_C}) \alpha k_1 M_{C,i}. \tag{31}$$

Rewriting equation 31 leads to the following expression for the lumped parameter:

$$\kappa_i = \phi_{C,i}^M \alpha k_i = \frac{\bar{\lambda}_i}{1 - \frac{\bar{\lambda}_i}{\lambda_C}} \tag{32}$$

Thus the value of κ_i can be calculated using only the growth law parameter λ_C and the steady-state growth rate on substrate i. This stands in contrast to Towbin *et al*, who had to fit κ_i for each substrate i separately. Instead, Erickson

et al derive a simple analytic expression.

Despite all this it is still not possible to simulate equation 29 because it is not possible to calculate the value of the factor $\phi_{C,i}^M \alpha k_j$, where $j \neq i$. We can solve this by instead simulating a differential equation for $\tilde{M}_{C,i} = \frac{M_{C,i}}{\phi_{C,i}^M}$ as follows:

$$\frac{\mathrm{d}\tilde{M}_{C,i}}{\mathrm{d}t} = \alpha(k_1 M_{C,1} + k_2 M_{C,2})\mathcal{X}_C = (\kappa_1 \tilde{M}_{C,1} + \kappa_2 \tilde{M}_{C,2})\mathcal{X}_C \qquad (33)$$

Note that the tilde symbol has a different interpretation than in the Chapter 2 above. While the normalised protein mass $\tilde{M}_{C,i}$ has no direct physical interpretation, it is clearly proportional to the mass of the C sector. Because the equations for the masses of the transport proteins are decoupled from the rest of the system we do not need to know the exact values of $M_{C,1}$ and $M_{C,2}$. Furthermore, in the plots in which the mass of the C sector is plotted (such as figure 6A and figure 7A) all protein masses are normalised, making the exact interpretation of $\tilde{M}_{C,i}$ irrelevant for the plot, as long as the quantity is proportional to the C sector.

We assume that the transporter of the substrate added at $t = 0$ is not produced in the absence of its substrate. Therefore, $\tilde{M}_{C,2} = 0$ and $\frac{\mathrm{d}\tilde{M}_{C,2}}{\mathrm{d}t} = 0$ until $t = 0$. From $t = 0$, $\frac{\mathrm{d}\tilde{M}_{C,2}}{\mathrm{d}t}$ is calculated using equation 33.

3.2.4 To simulate downshifts the external environment needs to be modelled

Because during downshifts the substrate that is being depleted is no longer present in saturating amounts, it is necessary to simulate n_2 in order to calculate the value of k_2 using equation 17. Erickson *et al* do not explicitly mention in their paper how they model the concentration of the nutrient. As a result we had to derive our own differential equation for the time evolution of n_2 in order to be able to simulate their model. The result is the following equation:

$$\frac{\mathrm{d}n_2}{\mathrm{d}t} = -\kappa_2 \tilde{M}_{C,2} \frac{n_2}{V_E(K_{Ms,2} + n_2)}. \qquad (34)$$

Here V_E is the volume of the growth medium. Equation 34 can be interpreted as saying that the substrate molecules are taken up with rate $\kappa_2 \frac{n_2}{K_{Ms,2}+n_2}$ by a transport protein with mass $M_{C,2}$. The factor α in κ_2 converts to units of protein mass. Therefore n_2 is in terms of units of protein mass per volume.

3.2.5 Simulating the dynamics of the translational activity

Erickson *et al* derive a differential equation for the translational activity from equation 23:

$$\frac{\mathrm{d}\sigma}{\mathrm{d}t} = \frac{\mathrm{d}\frac{J_R}{M_R}}{\mathrm{d}t} = \frac{1}{M_R}\frac{\mathrm{d}J_R}{\mathrm{d}t} - \frac{J_R}{M_R^2}\frac{\mathrm{d}M_R}{\mathrm{d}t} \tag{35}$$

.

During upshifts when k_2 is a constant we can simplify equation 35 by substituting equation 18, equation 16, equation 29 and equation 30:

$$\frac{\mathrm{d}\sigma}{\mathrm{d}t} = \sigma\left(\sum_{i=1}^{2}\alpha k_i\phi_{C,i}^{M}\mathcal{X}_C(\sigma) - \mathcal{X}_R(\sigma)\sigma\right) = \sigma\left((\kappa_1 + \kappa_2)\mathcal{X}_C(\sigma) - \mathcal{X}_R(\sigma)\sigma\right). \tag{36}$$

Note that this equation is only valid in the post-shift medium, because there is no uptake of substrate 2 before the shift. Also note that equation 36 is only a function of σ and thus the value of the central internal variable σ can be simulated separately from the rest of the system as long as the steady-state growth rate on both substrates are known (which are necessary to calculate the value of $\kappa_1 + \kappa_2$).

In subsection 3.2.1 we explained that during downshifts k_2 becomes a function of the concentration of the corresponding substrate. k_2 is now no longer constant in time. As a result we had to derive a new differential equation for σ to simulate downshifts. Erickson *et al* did not provide this derivation in their paper or supplement, but must have performed the same analysis. Equation 35 becomes

$$\frac{\mathrm{d}\sigma}{\mathrm{d}t} = \frac{1}{M_R}\frac{\mathrm{d}J_R}{\mathrm{d}t} + \alpha\frac{M_{C,2}}{M_R}\frac{\mathrm{d}k_2}{\mathrm{d}t} - \frac{J_R}{M_R^2}\frac{\mathrm{d}M_R}{\mathrm{d}t} \tag{37}$$

during downshifts. We can derive an equation for $\dfrac{\mathrm{d}k_2}{\mathrm{d}t}$ by differentiating equation 17. Using equation 34 and applying the quotient rule leads to

$$\frac{\mathrm{d}k_2}{\mathrm{d}t} = -k_{2,max}\frac{K_{Ms,2}\tilde{M}_{C,2}\mu_2 n_2}{V_E(K_{M,2} + n_2)^3}. \tag{38}$$

Now substituting equations 18, 16, 29, 30 and 38 into equation 37 gives rise to

$$\frac{\mathrm{d}\sigma}{\mathrm{d}t} = \sigma\left(\left(\kappa_1 + \kappa_2\frac{n_2}{K_{Ms,2} + n_2}\right)\mathcal{X}_C(\sigma) - \mathcal{X}_R(\sigma)\sigma\right) - \kappa_2^2\frac{\tilde{M}_{C,2}^2}{M_R}\frac{K_{Ms,2}n_2}{V_E(K_{Ms,2} + n_2)^3}. \tag{39}$$

Note that $\dfrac{\mathrm{d}\sigma}{\mathrm{d}t}$ is no longer a function of just σ. Nevertheless, equation 39 allows for the simulation of downshifts using the model from Erickson *et al.*

3.2.6 Downshifts require more parameters than those measured in growth laws

Because the regulation functions from Erickson *et al* are based on the growth laws the parameter values can be measured experimentally. Erickson *et al* perform these experiments themselves. The resulting parameter values are $\lambda_C = 1.17$, $\phi_{R,0} = 0.049$ and $\gamma = 11.02$. They also measure the value of $\tau = \frac{M_P}{M_B}$ and show this value to be approximately constant, such that in the simulations we set $\tau = 0.7$.

To simulate the downshifts we need the value of $K_{Ms,2}$. For the substrates used in the downshift experiments Erickson *et al* took this value from the literature. For the downshifts in which glucose is depleted the value of $K_{Ms,2} = 5$, for glycerol $K_{Ms,2} = 5.6$ and for gluconate $K_{Ms,2} = 212$.

To calculate the value of κ_i we need to know the steady-state growth rate on substrate i. Erickson *et al* also obtained these values through experiments. We will not list all the measured steady-state growth rates here, but they can be found on page 20 of the supplement to their paper. The co-utilized growth rate $\bar{\lambda}_{1,2}$ (on a medium with both substrates present in saturating concentrations) can be calculated approximately from $\bar{\lambda}_1$ and $\bar{\lambda}_2$ using the composition formula [16] as follows:

$$\lambda_{1,2} = \frac{\lambda_1 + \lambda_2 + 2\frac{\lambda_1 \lambda_2}{\lambda_C}}{1 - \frac{\lambda_1 \lambda_2}{\lambda_C^2}}. \tag{40}$$

Note that all parameters have been measured or can be calculated from the experimental measurements done by Erickson *et al*. In fact the model is entirely parameter free when modelling upshifts. Not a single parameter need to be fit, and all the dynamics follow from the growth laws. This is an astounding feature of the model.

However, when when modelling downshifts an initial conditions for the external environment has to be set, due to the fact that substrate 2 is no longer present in saturating amounts at the time of the downshift. To be able to calculate equation 17 one now has to simulate n_2. Simulating this nutrient requires setting an initial condition. We do this by specifying both an initial condition for the number of substrate particles $N_{2,in}$ and the volume of the growth medium V_E. We then calculate $n_{2,in}$ according to

$$n_{2,in} = \frac{N_{2,in}}{V_E}. \tag{41}$$

We have to specify $N_{2,in}$ and V_E separately because the value of V_E is also required in equation 34. We choose the values of $N_{2,in}$ and V_E so as to best match the model to the data (by eye), having no further information on these parameters. We applied the Erickson model to the same data as in figure 7 while

varying the parameter $N_{2,in}$ (Figure 9A) and the parameter V_E (Figure 9B) to investigate the effect on the dynamics of the growth rate during a downshift. The effect of the $N_{2,in}$ parameter can be loosely described as setting the approximate time at which the growth rate starts to decrease. By tweaking V_E we can control how gradually the growth rate decreases. If the growth rate decreases in a more gradual manner it dips less below its eventual steady-state value.
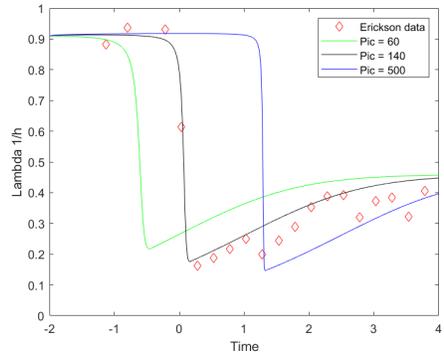
(a)
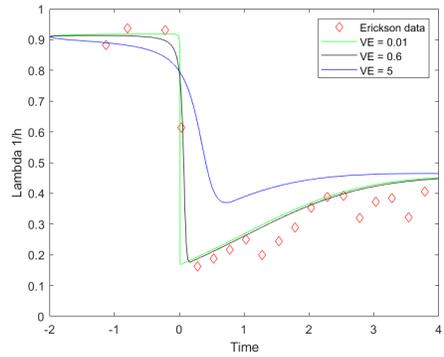
(b)
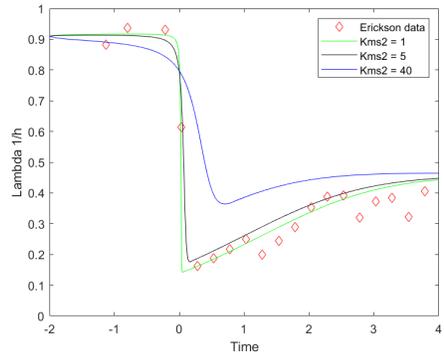
(c)

Figure 9: Hidden parameters effect the dynamics of downshifts. Initial growth is on succinate and glucose. Around t = 0 glucose is depleted. Parameters are (unless specified in plot legend): $\bar{\lambda}_1 = 0.46$, $\bar{\lambda}_2 = 0.88$, $K_{Ms,2} = 5$, $V_E = 0.6$, $N_{2,in} = 140$. Time is in hours. **a** Varying the $N_{2,in}$ parameter changes the time at which the downshift occurs. **b** Varying the V_E parameter changes how gradually the downshift occurs. **c** Varying the K_{Ms2} parameter has a similar effect to varying V_E.

29

The effect of varying the V_E parameter is similar to varying the Michaelis Menten parameter of the substrate (Figure 9C). Despite the fact that this parameter is derived from the literature and has a fixed value, the substrates can be approximately modelled as having a different value for $K_{Ms,2}$ by varying the value of V_E.

Because Erickson *et al* do not derive an equation for the dynamics of nutrient depletion, they also make no mention of the parameters $N_{2,in}$ and V_E. Despite this, these parameters are an indispensable part of the Erickson *et al* model. And while their model can match the data in the upshift experiments by merely specifying $\bar{\lambda}_1$ and $\bar{\lambda}_2$, the hidden parameters $N_{2,in}$ and V_E allow for much more freedom when modelling downshifts, assuming the authors choose these values to best fit the data. This seems to me the most likely scenario. Of course, the actual values of these parameters are likely known by the authors. Nevertheless, it seems probable that if Erickson *et al* had been able to fit the downshifts with known values for these parameters they would have mentioned it. However, these parameters are not mentioned in either the main paper or the supplement.

3.3 We have reproduced the match of the Erickson *et al* model to the up- and downshift experiments

We have reproduced the results obtained by Erickson *et al* in order to confirm them. This has also allowed us to gain additional insight into their system, since we were able to visualise the time evolution of variables which were not plotted in the original paper. In figure 10A, B and C we plot our reproduction of the results by Erickson *et al* that we showed in figure 6. We can thus validate that the Erickson *et al* model provides a good match to experimental data of the upshift from succinate to succinate and gluconate. Let us reiterate that all parameter values, being either constants from the growth laws or steady-state growth rates, follow directly from experiment. These entirely determine the dynamics of the system. In the supplementary materials we show the fit to the up- and downshift experiments with succinate and glycerol.

(a)

(b)

(c)

(d)

(e)

(f)

(g)
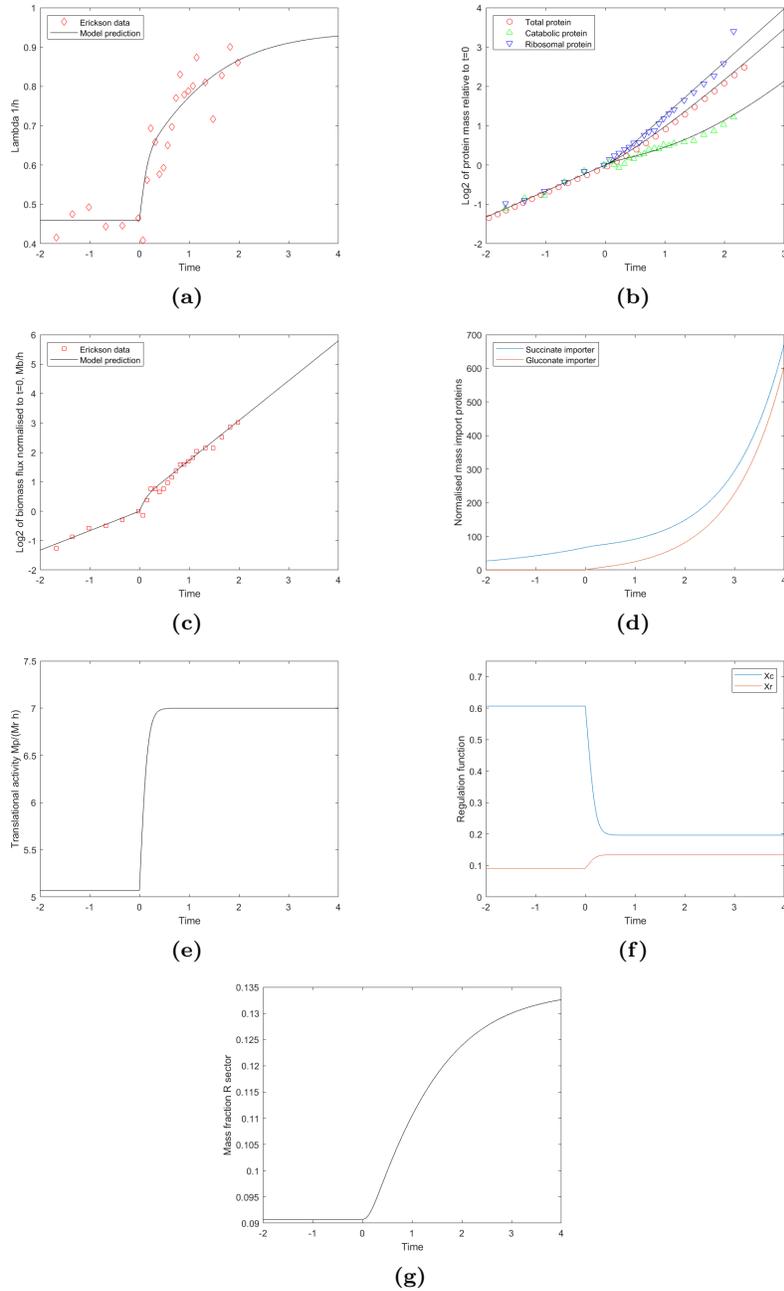
Figure 10: The Erickson *et al* model quantitatively matches upshift experiment. Initial growth is on succinate. At $t = 0$ gluconate is added. Data from [11]. Parameters used: $\bar{\lambda}_1 = 0.46$, $\bar{\lambda}_2 = 0.91$. Dots are experimental measurement, lines are model predictions. Time is in hours. **a** Growth rate. **b** Protein masses. **c** Biomass flux. **d** $\tilde{M}_{C,1}$ and $\tilde{M}_{C,2}$. **e** Translational activity. **f** Regulation functions. **g** Mass fraction R sector

31

Figure 10D shows that the mass of the protein responsible for transporting gluconate starts to increase from $t = 0$. As a result, the protein flux per unit of protein mass increases (figure 10A). This in turn causes the translational activity to increase (figure 10E). Because $\mathcal{X}_R$ is a positive function of the translational activity (equation 26) the cell allocates a larger part of its proteome towards the R sector (figure 10F) and the mass fraction of the R sector starts to increase (figure 10G). Because $\lambda = \frac{J_B}{M_B} = \frac{J_R}{M_P}$ increases by a larger factor than the mass fraction of the R sector, the translational activity (which is equal to $\frac{\lambda}{\phi_R}$) reaches a steady-state higher than during growth on succinate only. This results in lower C sector allocation and higher R sector allocation. All intensive variables reach their respective steady-state values monotonically.

Figure 11A, B and C shows our reproduction of the glucose downshift depicted in figure 7. By choosing $N_{2,in}$ and V_E by eye, the Erickson *et al* model matches the data quite accurately. This demonstrates that the differential equation for the nutrient (equation 34) and the differential equation for σ (equation 39) we derived earlier successfully capture the dynamics of nutrient depletion.
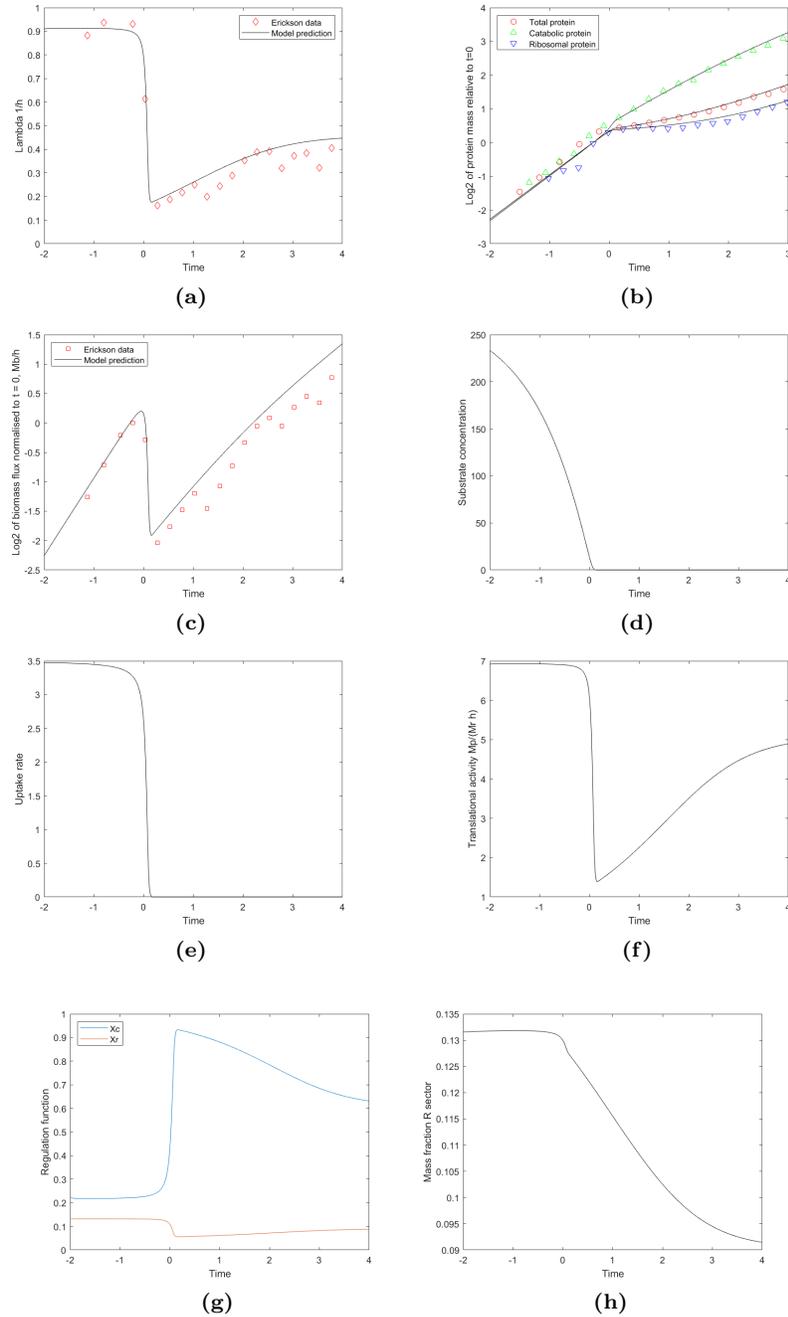
**Figure 11:** The Erickson *et al* model quantitatively matches downshift experiment. Initial growth is on succinate and glucose. Around t = 0 glucose is depleted. Data from [11]. Parameters used: $\bar{\lambda}_1 = 0.46$, $\bar{\lambda}_2 = 0.88$, $N_{2,in} = 140$, $V_E = 0.6$. Dots are experimental measurement, lines are model predictions. Time is in hours. **a** Growth rate. **b** Protein masses. **c** Biomass flux. **d** Concentration glucose. **e** Effective uptake parameter glucose importer. **f** Translational activity **g** Regulation functions. **h** Mass fraction R sector.

Figure 11D shows that the nutrient is being depleted during the initial phase of the simulation. As a result, the effective uptake parameter (equation 17) decreases (figure 11E) before rapidly approaching zero. As there is suddenly no flux through the glucose transporter, the protein flux per unit of protein mass (figure 11)A rapidly drops, and as a consequence so does the translational activity (figure 11)F. This causes E. coli to allocate less of its protein budget to the R sector, and more to the C sector (figure 11G). As a result the mass fraction of the R sector starts to slowly approach its post-shift steady-state value, causing the translational activity to recover and approach its new steady-state value. The growth rate shows a similar dynamic to the translational activity, first displaying a dip and subsequently a partial recovery. As the nutrient is depleted there is suddenly not enough of the glucose transport protein to supply the ribosomes with enough resources. As the C sector is up-regulated and the R sector is down-regulated the growth rate increases.

3.4   Why do we need the ErTo model?

As we have shown, the model by Erickson *et al* is capable of modelling dynamical data. The fact that the Erickson *et al* model works shows that what we know about steady-state proteome allocation (i.e. the growth laws) is also sufficient to explain how E. coli transitions between steady-states. After all, the regulation functions were derived using only steady-state relationships. Observe how, like Towbin *et al*, this model once again has only a single internal variable (σ)).

One might be inclined to wonder, after having read about two such successful models, why another model is needed. The model of Erickson *et al* is incapable of describing the O-curves of the mutant used by Towbin *et al*. This is because Erickson *et al* do not model product inhibition, meaning their model is incapable of over expressing the C sector. Vice versa, we cannot apply the Towbin *et al* model to the data of Erickson *et al*, because their model requires the fitting of substrate specific parameters to the O-curve data. This is an obstacle because in the experiments of the two studies disparate carbon substrates were used. On top of that, the model of Towbin *et al* is a single substrate model, and as such unsuitable for modelling up- and downshifts.

Due to these considerations we present our own model of protein allocation in *E. coli*. It can be applied to the data of both Towbin *et al* and Erickson *et al*.

4 A hybrid model can describe both O-curve and dynamical data

In this chapter, we will derive a two substrate model with no substrate specific fitting parameters. This model can be applied to both the wildtype and mutant steady-state data from Towbin *et al*, and the dynamical data from Erickson *et al*.

4.1 Deriving the single substrate version of the ErTo model

4.1.1 We follow Towbin *et al* in coarse graining the proteome

Like Towbin *et al*, we model the proteome as consisting of three protein sectors: The C sector, the R sector and the Q sector. Consequently, their respective mass fractions must sum to one:

$$\phi_C + \phi_R + \phi_Q = 1. \tag{42}$$

We treat the Q sector as maintaining a constant fraction of the proteome.

4.1.2 The fluxes are a function of the internal metabolite

In line with Towbin *et al*, we assume the cell senses its environment via the concentration of an internal metabolite x which is central between catabolism and translation. We assume that transport of the carbon substrate by its specific transport protein is subject to product inhibition by x and that the rate of protein synthesis saturates with increasing x. Therefore, the carbon flux J_C and the ribosomal flux J_R on substrate i, both in units of protein mass per time, are given by the following equations:

$$J_C = \alpha \frac{k_{i,max} M_{C,i} K_I}{K_I + x} \tag{43}$$

and

$$J_R = \frac{\gamma M_R x}{K_M + x}. \tag{44}$$

Here α converts from carbon flux in terms of carbon atoms to units of protein mass, $k_{i,max}$ is the maximal uptake rate in terms of carbon atoms of an transport protein with mass $M_{C,i}$, K_I is the value of x at which the transport of substrate i is half maximal, K_M is the value of x at which the rate of translation is half maximal and γ denotes the maximal rate of protein production by ribosomal proteins with mass M_R. In equation 43 we have assumed that substrate i is present in saturating amounts.

We assume that the mass of any transport protein is always proportional to the mass of the C-sector as a whole:

$$M_{C,i} = \frac{\phi_{C,i}^M}{1 - \phi_Q} M_C. \tag{45}$$

This assumption is true for most carbon substrates but not for some like glycerol or pyruvate, both of which have a baseline expression independent of the expression of the C sector [10]. In equation 45 $\phi_{C,i}^M$ denotes the proportionality constant which scales down $M_{C,i}$ from $\frac{M_C}{1-\phi_Q}$ (or $\phi_{C,i}$ from $\frac{\phi_C}{1-\phi_Q}$). Because we do not know the exact values of the parameters α, $k_{i,max}$ or $\phi_{C_i}^M$, we lump them together in a single parameter like Erickson *et al*:

$$\kappa_i = \alpha k_{i,max} \frac{\phi_{C,1}^M}{1 - \phi_Q}. \tag{46}$$

Here κ_i can be interpreted as the maximal uptake rate in terms of carbon atoms on a particular substrate. This quantity therefore represents growth potential or quality of the environment. Using this notation, equation 43 now becomes

$$J_C = \frac{\kappa_i M_C K_I}{K_I + x}. \tag{47}$$

Like Towbin *et al* we define the growth rate to be equal (up to a constant) to the protein flux per unit of protein mass:

$$\lambda(x, \phi_R) = \frac{1}{M_P} \frac{\mathrm{d}M_P}{\mathrm{d}t} = \frac{1}{M_P} J_R. \tag{48}$$

4.1.3 Deriving the equation from Towbin *et al* for the central metabolite

We can now postulate a differential equation for X, which describes the change in the amount of X as being the result of a balance between the carbon and the ribosomal flux:

$$\frac{\mathrm{d}X}{\mathrm{d}t} = \rho(\alpha J_C - J_R). \tag{49}$$

Here ρ is a factor that converts from protein mass to amount of metabolic intermediate. However, we want an equation for the metabolic intermediate as an intensive variable, denoted by x, in order to be able to model product inhibition and ribosome saturation. We can derive this from equation 49 resulting in

$$\frac{\mathrm{d}\frac{X}{\rho M_P}}{\mathrm{d}t} = \frac{\mathrm{d}x}{\mathrm{d}t} = \alpha \frac{J_C}{M_P} - \frac{J_R}{M_P} - \lambda x = \frac{\kappa_i \phi_C K_I}{K_I + x} - \frac{\gamma \phi_R x}{K_M + x} - \lambda x. \tag{50}$$

The units of x are now $\frac{X}{\rho M_P}$. The factor $\frac{1}{\rho}$ converts from amount of X to the amount of protein mass which could be produced from this, and the factor $\frac{1}{M_P}$ converts to a concentration (since M_P is proportional to the volume of the cells).

Alternatively, ρM_P can be seen as the amount of X needed to produce M_P.

We assume that the dilution of x is negligible with respect to the rate of uptake/translation and scale x with a factor $\frac{1}{K_M}$ to remove a parameter that we would have to fit later. Note that this also changes the interpretation of γ and κ_1. Equation 50 now becomes

$$\frac{\mathrm{d}\frac{x}{K_M}}{\mathrm{d}t} = \frac{\mathrm{d}\hat{x}}{\mathrm{d}t} = \frac{\kappa}{K_M}\phi_C \frac{\frac{K_I}{K_M}}{\frac{K_I}{K_M} + \frac{x}{K_M}} - \left(\frac{\gamma}{K_M}\right)\frac{\frac{x}{K_M}}{1 + \frac{x}{K_M}}\phi_R$$
$$= \frac{\hat{\kappa}_i \phi_C \hat{K}_I}{\hat{K}_I + \hat{x}} - \frac{\hat{\gamma}\phi_R \hat{x}}{1 + \hat{x}}, \tag{51}$$

the central equation of our model. Variables carrying a circumflex are scaled by a factor $\frac{1}{K_M}$. Note that equation 51 is identical to equation 4 used by Towbin *et al* to model $\frac{\mathrm{d}\hat{x}}{\mathrm{d}t}$, while all parameters carry similar interpretations as those in Erickson *et al.*

We assume that the dynamics of x will be much faster than the dynamics of protein allocation such that the carbon- and protein fluxes are always in flux balance on the timescale of protein production. This makes sense because a single protein consists of many units of x. We write:

$$\frac{\mathrm{d}\hat{x}}{\mathrm{d}t} = 0. \tag{52}$$

Note that equation 52 implies that we could also have defined the growth rate on substrate i through the normalised carbon flux, $\alpha \frac{\hat{\kappa}_i \phi_C \hat{K}_I}{\hat{K}_I + \hat{x}}$. After all, the quasi-steady-state assumption of x means that the carbon- and protein flux are always in flux balance. We solve equation 52 for $\hat{x}$ to obtain the quasi-steady-state of $\hat{x}$ as a function of ϕ_C and ϕ_R, denoted by $\hat{\bar{x}}(\phi_C, \phi_R)$, which results in the following equation:

$$\hat{\bar{x}}(\phi_C, \phi_R) = \frac{\hat{K}_I(\hat{\kappa}_i\phi_C + \hat{\gamma}\phi_R) + \sqrt{4\hat{K}_I\hat{\gamma}\hat{\kappa}_i\phi_C\phi_R + (\hat{K}_I(\hat{\kappa}_i\phi_C - \hat{\gamma}\phi_R))^2}}{2\hat{\gamma}\phi_R}. \tag{53}$$

From equations 48 and 53 we can now derive an expression for the growth rate in terms of ϕ_C and ϕ_R as

$$\lambda(\phi_C, \phi_R) = \frac{\hat{\gamma}\phi_R \hat{\bar{x}}(\phi_C, \phi_R)}{1 + \hat{\bar{x}}(\phi_C, \phi_R)} = \frac{\hat{K}_I(\hat{\kappa}_i\phi_C + \hat{\gamma}\phi_R) - \sqrt{(\hat{K}_I\hat{\kappa}_i\phi_C - \hat{K}_I\hat{\gamma}\phi_R)^2 + 4\hat{K}_I\hat{\gamma}\hat{\kappa}_i\phi_C\phi_R}}{2(\hat{K}_I - 1)}. \tag{54}$$

4.1.4 The regulation functions of ErTo model are derived by assuming optimality

It is in the derivation of the regulation functions that we take a different approach from either Towbin *et al* or Erickson *et al*. Rather than using the growth laws as an empirical basis or fitting a flexible function we assume that E. coli regulates optimally (i.e. maximises the growth rate on whatever substrate it is growing on). Of course, we have seen in chapter 2 that regulation does not always maximise the growth rate. Our model will be unsuitable for substrates for which this is the case. While Towbin *et al* show for which carbon substrates regulation is optimal, they use mostly different substrates from Erickson *et al*. For most carbon substrates used in the dynamical experiments we do not know whether they are optimally regulated. Nevertheless, Towbin *et al* found that more than half of the substrates they looked at maximised the growth rate. If we assume this is also true for the substrates used by Erickson *et al*, we expect to be able to model at least some of their up- and downshifts.

To calculate the optimal ϕ_C in a given environment, denoted by ϕ_C^*, we solve

$$\left.\frac{\mathrm{d}\lambda(\phi_C, 1 - \phi_Q - \phi_C)}{\mathrm{d}\phi_C}\right|_{\phi_C=\phi_C^*} = 0 \tag{55}$$

for ϕ_C. This results in the following equations for $\phi_C^*(\hat{\kappa}_i)$ and $\phi_R^*(\hat{\kappa}_i)$, the optimal concentrations of the protein sectors in a given environment:

$$\phi_C^*(\hat{\kappa}_i) = (1 - \phi_Q)\frac{\hat{\kappa}_i\sqrt{\hat{K}_I\gamma\hat{\kappa}_i} + \hat{K}_I\gamma(\gamma + \hat{\kappa}_i) - \gamma(2\hat{\kappa}_i + \sqrt{\hat{K}_I\gamma\hat{\kappa}_i})}{\hat{K}_I(\gamma + \hat{\kappa}_i)^2 - 4\gamma\hat{\kappa}_i} \tag{56}$$

and

$$\phi_R^*(\hat{\kappa}_i) = 1 - \phi_Q - \phi_C^*(\hat{\kappa}_i) = (1 - \phi_Q)\frac{\hat{K}_I\hat{\kappa}_i(\gamma + \hat{\kappa}_i) + \gamma\sqrt{\hat{K}_I\gamma\hat{\kappa}_i} - 2\gamma\hat{\kappa}_i - \hat{\kappa}_i\sqrt{\hat{K}_I\gamma\hat{\kappa}_i}}{\hat{K}_I(\gamma + \hat{\kappa}_i)^2 - 4\gamma\hat{\kappa}_i}. \tag{57}$$

We postulate that the cell infers its environment from the concentration of the internal metabolite x. That is, we want to obtain an expression for $\phi_C^*(\hat{x})$ and $\phi_R^*(\hat{x})$. In order to do this, we first solve the following equation for $\hat{\kappa}_i$:

$$\hat{\bar{x}}(\phi_C^*(\hat{\kappa}_i)), \phi_R^*(\hat{\kappa}_i)) = \hat{\bar{x}}. \tag{58}$$

This leads to:

$$\hat{\kappa}_1(\hat{\bar{x}}) = \frac{\hat{\bar{x}}^2\gamma}{\hat{K}_I}. \tag{59}$$

By substituting equation 59 into equations 56 and 57 we can now derive the optimal regulation functions that determine what fraction of the protein flux

is allocated to the C- or R sector respectively. This results in the following equations:

$$\mathcal{X}_C(\hat{x}) = \phi_C^*(\hat{\kappa}_1(\hat{x})) = (1 - \phi_Q)\frac{\hat{x} + \hat{K}_I}{\hat{x}(2 + \hat{x}) + \hat{K}_I} \tag{60}$$

and

$$\mathcal{X}_R(\hat{x}) = \phi_R^*(\hat{\kappa}_1(\hat{x})) = (1 - \phi_Q)\frac{\hat{x}(1 + \hat{x})}{\hat{x}(2 + \hat{x}) + \hat{K}_I}. \tag{61}$$

Note that our regulation functions depend on the parameter $\hat{K}_I$. Therefore, the optimal regulation functions do not allow us to make a substrate specific fit of the product inhibition constant $\hat{k}_I$. The pathway for the catabolism of one substrate might have an entirely different $\hat{k}_I$ than another. Values could also be very similar There does not appear to be any data on this as of yet. We have taken a bit of a gamble with this assumption, for the reason that this allows our model to describe the Erickson *et al* using a fit on the Towbin *et al* data. After all, with a global value for $\hat{K}_I$ our model does not have to fit any substrate specific parameters.

The quantities $\frac{\mathcal{X}_C}{1 - \phi_Q}$, $\frac{\mathcal{X}_C}{1 - \phi_Q}$ and their sum are plotted in figure 12.
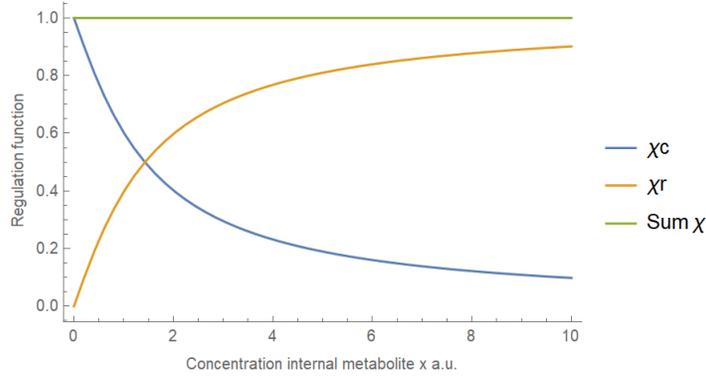

Figure 12: Regulation functions of the ErTo model scaled between 0 and 1.

4.1.5 Dynamics of protein sectors follow from optimal allocation of protein flux and dilution

We model protein flux as being optimally distributed between the C- and R sectors. This results in the following differential equations for the masses of the C- and R sectors:

$$\frac{\mathrm{d}M_C}{\mathrm{d}t} = \mathcal{X}_C J_R \tag{62}$$

and

$$\frac{\mathrm{d}M_R}{\mathrm{d}t} = \mathcal{X}_R J_R \tag{63}$$

From equations 62 and 63 we can derive equations for $\dfrac{\mathrm{d}\phi_C}{\mathrm{d}t}$ and $\dfrac{\mathrm{d}\phi_R}{\mathrm{d}t}$ using the quotient rule, as below:

$$\frac{\mathrm{d}\phi_C}{\mathrm{d}t} = \frac{\mathrm{d}\frac{M_C}{M_P}}{\mathrm{d}t} = \lambda(\mathcal{X}_C - \phi_C) \tag{64}$$

and

$$\frac{\mathrm{d}\phi_R}{\mathrm{d}t} = \frac{\mathrm{d}\frac{M_R}{M_P}}{\mathrm{d}t} = \lambda(\mathcal{X}_R - \phi_R). \tag{65}$$

4.1.6 Fitting the ErTo model to the Towbin *et al* data without substrate specific parameters

We fit our model to the c-line and O-curve data from Towbin *et al*. Because we assume optimality we only fit to the O-curves of substrates which Towbin *et al* demonstrated to be optimally regulated. These substrates are lactose (with the addition of TDG, a lactose permease inhibitor, at various concentrations), sorbitol and maltose. We multiply all growth rates measured by Towbin *et al* by a factor of 1.6. This is because the study done by Towbin *et al* measures lower growth rates than those of the in the Hwa lab (where the Erickson *et al* experiments were performed). This could be because Towbin *et al* performed their experiments on plate wells, whereas Erickson *et al* used a batch culture.

We derive the c-line from our model by solving

$$\phi_C^*(\hat{\kappa}) = \phi_C \tag{66}$$

for $\hat{\kappa}$. In this way we obtain an equation for $\hat{\kappa}(\phi_C^*)$, the perceived environment in steady state with optimal ϕ_C. The c-line is defined by $\lambda(\phi_C, 1 - \phi_Q - \phi_C)[\hat{\kappa}(\phi_C^*)/\kappa_i]$. Here we have derived an expression for the growth rate as a function of the optimal ϕ_C. Because in our model ϕ_C is optimal in steady-state, this relationship constitutes the c-line. Observe that the c-line we derive is rather like the one derived by Towbin *et al* (figure 13), in that it is only approximately linear for high enough values of $\bar{\lambda}$.

To derive the O-curves we solve

$$\bar{\lambda} = \lambda(\phi_C, 1 - \phi_Q - \phi_C)[\phi_C^*(\hat{\kappa}_i)/\phi_C] \tag{67}$$

for $\hat{\kappa}_i$. The resulting expression for the value of the lumped parameter $\hat{\kappa}_i$ as a function of the steady-state growth rate is

$$\hat{\kappa}_i(\bar{\lambda}) = \frac{\hat{\gamma}\bar{\lambda}(2(\bar{\lambda} + \sqrt{\bar{\lambda}(\bar{\lambda} - \hat{K}_I(\bar{\lambda} + \hat{\gamma}(\phi_Q - 1)))}) - \hat{K}_I(\bar{\lambda} + \hat{\gamma}(\phi_Q - 1)))}{\hat{K}_I(\bar{\lambda} + \hat{\gamma}(\phi_Q - 1))^2}. \quad (68)$$

Observe that, like Erickson et al we now have a function for $\hat{\kappa}_i$ as a function of $\bar{\lambda}$, although we derived the relationship from an optimality principle rather than the growth laws. Equation 68 means we will not have to fit the value of $\hat{\kappa}_i$ to the O-curve data on substrate i. The model expression of the O-curve on substrate i with steady-state growth rate $\bar{\lambda}$ is defined by $\lambda(\phi_C, 1 - \phi_Q - \phi_C)[\hat{\kappa}_i(\bar{\lambda})/\hat{\kappa}_i]$. By choosing the value of $\hat{\kappa}_i$ which produces the correct $\bar{\lambda}$ at the optimal ϕ_C, we ensure that the top of the O-curve matches with the measured steady-state growth rate. The rest of the functional form of the O-curve follows from the sum constraint and equation 53. The steady-state growth rates were measured by Towbin et al.

Apart from the model parameters we also fit the factor θ which converts from the CRP activity measured by Towbin et al to ϕ_C. We simultaneously fit the c-line and O-curves to the Towbin et al data under the constraint that $0.34 < \phi_Q < 0.7$ and $\gamma(1 - \phi_Q) = 1.2$. The latter quantity corresponds to the maximal steady-state growth rate (equivalent to λ_C in the model from Erickson et al). The resulting parameter values (which we have used in all results presented in this chapter) are as follows: $\hat{\gamma} = 2.72092$, $\hat{K}_I = 2.03501$, $\phi_Q = 0.558972$ and $\theta = 0.354854$.

Figure shows our fit to the c-line data from Towbin et al. In figure 14 we show the fit to the lactose O-curves. Our model can fit both types of data quite well.
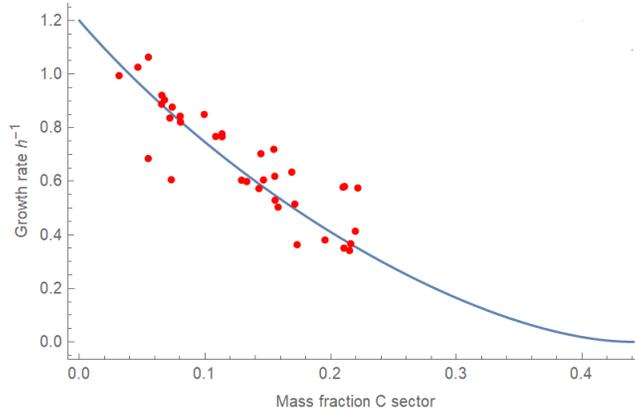

Figure 13: The c-line derived from the ErTo model is also a non-linear function. Blue line is the fit of the ErTo model to the data from Towbin et al (red dots).
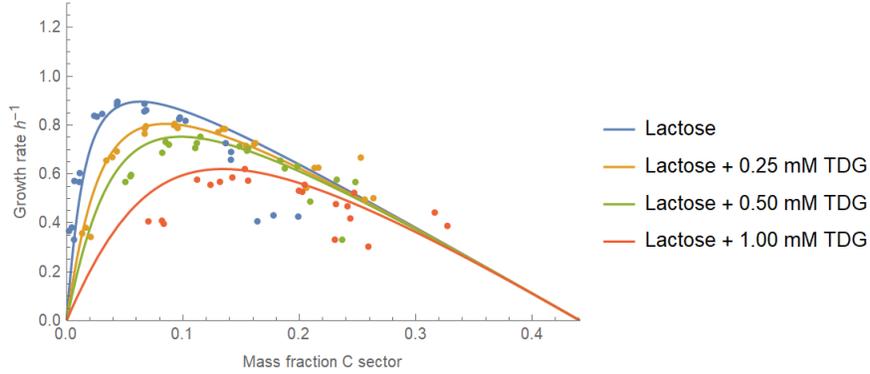
Figure 14: The Erto model provides a good fit to the O-curves on lactose. Solid lines are the fit of the ErTo model and dots are O-curve data from Towbin *et al*. E. coli was grown on lactose under the additions of varying amounts of the lactose transport inhibitor TDG.

Note that we have fitted all parameters globally, whereas Towbin *et al* fit the value of $\hat{K}_I$ and $\hat{\kappa}_i$ separately for each substrate. We instead assume that $\hat{K}_I$ is the same or at least similar for all transport proteins. We can calculate $\hat{\kappa}_i$ using equation 68, so we do not have to fit this parameter.

Using the fitted values of the global parameters we can compare the regulation functions of the Towbin et al and ErTo models (Figure 15). Our regulation functions have ended up being quite similar, and are virtually identical for low values of $\hat{x}$. Given that we also use similar equations for $\dfrac{\mathrm{d}\hat{x}}{\mathrm{d}t}$ and $\dfrac{\mathrm{d}\phi_C}{\mathrm{d}t}$ the steady-state predictions of the ErTo model and the predictions of the Towbin *et al* model must be quite similar.
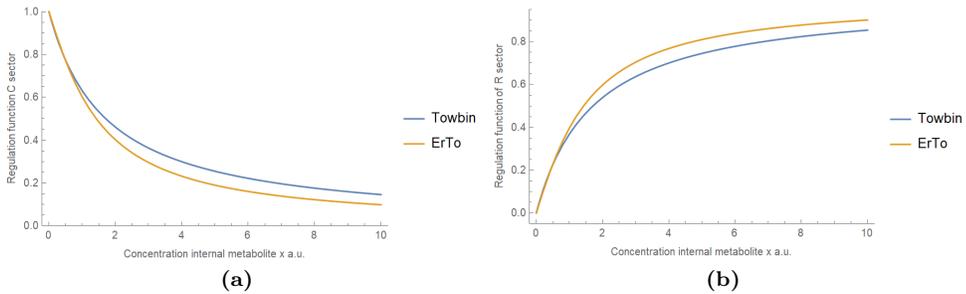

Figure 15: The regulation functions of the ErTo model and Towbin *et al* model are very similar. ErTo regulation functions are scaled with $\frac{1}{1-\phi_Q}$. **a** Regulation function C sector. **b** Regulation function R sector.

4.2 Extending the model to a two substrate model

Let $\frac{1-\phi_Q}{\phi_{C,i}^M} M_{C,i} = \tilde{M}_{C,i}$. Let the substrate which is present for the entire experiment be denoted as substrate 1, and let the substrate which is added at $t = 0$ be referred to as substrate 2. Note that $\frac{1-\phi_Q}{\phi_{C,1}^M} M_{C,1} = \tilde{M}_{C,1} = M_C$ (equation 45). We follow Erickson in modelling the transporter of the substrate which is added at $t = 0$ as absent until its substrate is added. Therefore, the quantity $\frac{1-\phi_Q}{\phi_{C,2}^M} M_{C,2} = \tilde{M}_{C,2}$ will only be equal to M_C in the steady-state after the upshift. We define $\tilde{\phi}_{C,i} = \frac{\tilde{M}_{C,i}}{M_P}$. We model $\tilde{\phi}_{C,2}$ as

$$\frac{\mathrm{d}\tilde{\phi}_{C,2}}{\mathrm{d}t} = \begin{cases} 0 & \text{if } t < 0 \\ \lambda(\mathcal{X}_C - \tilde{\phi}_{C,2}) & \text{if } t \geq 0. \end{cases} \tag{69}$$

This means $\tilde{M}_{C,2} = 0$ until $t = 0$ and after the upshift $\tilde{M}_{C,2} = M_C$. In equation 69 we assume that the sudden production of the second transport protein does not directly increase the mass fraction of the C-sector as a whole. That is, we assume $\phi_{C,2} << \phi_C$.

During upshift experiments

$$J_C = \sum_i^{n=2} \frac{\hat{\kappa}_i \tilde{M}_{C,i} K_I}{K_I + \hat{x}}. \tag{70}$$

Equation 51 then becomes

$$\frac{\mathrm{d}\hat{x}}{\mathrm{d}t} = \sum_i^{n=2} \frac{\hat{\kappa}_i \tilde{\phi}_{C,i} \hat{K}_I}{\hat{K}_I + \hat{x}} - \frac{\hat{\gamma}\phi_R \hat{x}}{1 + \hat{x}}. \tag{71}$$

From which the equation 71 the quasi-steady-state of $\hat{x}$ follows:

$$\hat{\bar{x}}(\tilde{\phi}_{C,1}, \tilde{\phi}_{C,2}, \phi_R) = \frac{\hat{K}_I(\sum_i \left(\hat{\kappa}_i \tilde{\phi}_{C,i}\right) + \hat{\gamma}\phi_R) + \sqrt{4\hat{K}_I \hat{\gamma} \sum_i \left(\hat{\kappa}_i \tilde{\phi}_{C,i}\right)\phi_R + (\hat{K}_I(\sum_i \left(\hat{\kappa}_i \tilde{\phi}_{C,i}\right) - \hat{\gamma}\phi_R))^2}}{2\hat{\gamma}\phi_R}. \tag{72}$$

If we want to simulate the system we need initial conditions for the concentrations of the protein sectors. By substituting equation 68 into equations 56 and 57 we can calculate $\phi_C^*(\bar{\lambda})$ and $\phi_R^*(\bar{\lambda})$. In this manner we can use the steady-state growth rates measured by Erickson *et al* to calculate the pre shift steady-state values of ϕ_C and ϕ_R.

4.3 Simulating the external environment

We model the uptake parameter of the substrate which is depleted around $t = 0$ as Erickson *et al* do:

$$\hat{k}_2 = \hat{k}_{2,max} \frac{n_2}{K_{Ms,2} + n_2}. \tag{73}$$

Here $\hat{k}_2$ is the uptake rate of the transport protein of the substrate which is being depleted and K_{Ms2} is the value of n_2 at which $\hat{k}_2$ is half maximal. We use the same values for K_{Ms2} as Erickson *et al*. To calculate the value of $\hat{k}_2$, we now need to explicitly model the concentration of the substrate. The depletion of the amount of substrate N_2 by its transporter of mass $M_{C,2}$ is governed by the differential equation below:

$$\frac{dN_2}{dt} = -\zeta \alpha J_{C,2} = -\zeta \hat{k}_2 M_C \frac{K_I}{K_I + \hat{x}} \frac{n_2}{K_{Ms,2} + n_2}. \tag{74}$$

Here ζ is a constant which converts from units of protein mass to number of substrate molecules and $J_{C,2}$ is the carbon flux through the transport protein which is being depleted. We can now derive an equation for the concentration of the substrate:

$$\frac{d \frac{N_2}{V_E}}{dt} = \frac{dn_2}{dt} = -\frac{\zeta \alpha J_{C,2}}{V_E} = -\frac{\zeta \hat{k}_2}{V_E} M_C \frac{K_I}{K_I + \hat{x}} \frac{n_2}{K_{Ms,2} + n_2}. \tag{75}$$

Here V_E is the volume of the growth medium, which we choose to fit the data as well as possible. By specifying an appropriate initial condition for the number of substrate molecules $N_{2,in}$, we can set an initial condition for n_2 (using equation 41) which causes the substrate to be depleted around $t = 0$.

We use the same equations for $\frac{d\hat{x}}{dt}$ and the quasi-steady-state of $\hat{x}$ as during upshifts (equations 71 and 72).

When simulating the system we need to specify initial conditions for ϕ_C and ϕ_R. We obtain these by simulating the system on the two substrate medium (with both present in saturating concentrations) until steady-state. We then use the steady-state ϕ_C and ϕ_R as initial conditions for the downshift simulation.

4.4 The ErTo model has numerous limitations when modelling the upshift- and downshift experiments

Having extended the ErTo model to accommodate modelling upshifts and downshifts, we sought to apply it to model the experiments performed by Erickson *et al*. Figure 16 shows the match of the ErTo model to the gluconate upshift experiment. The match to the data is about as good as the fit of the Erickson *et al* model.

(a)

(b)

(c)

(d)

(e)

(f)

(g)
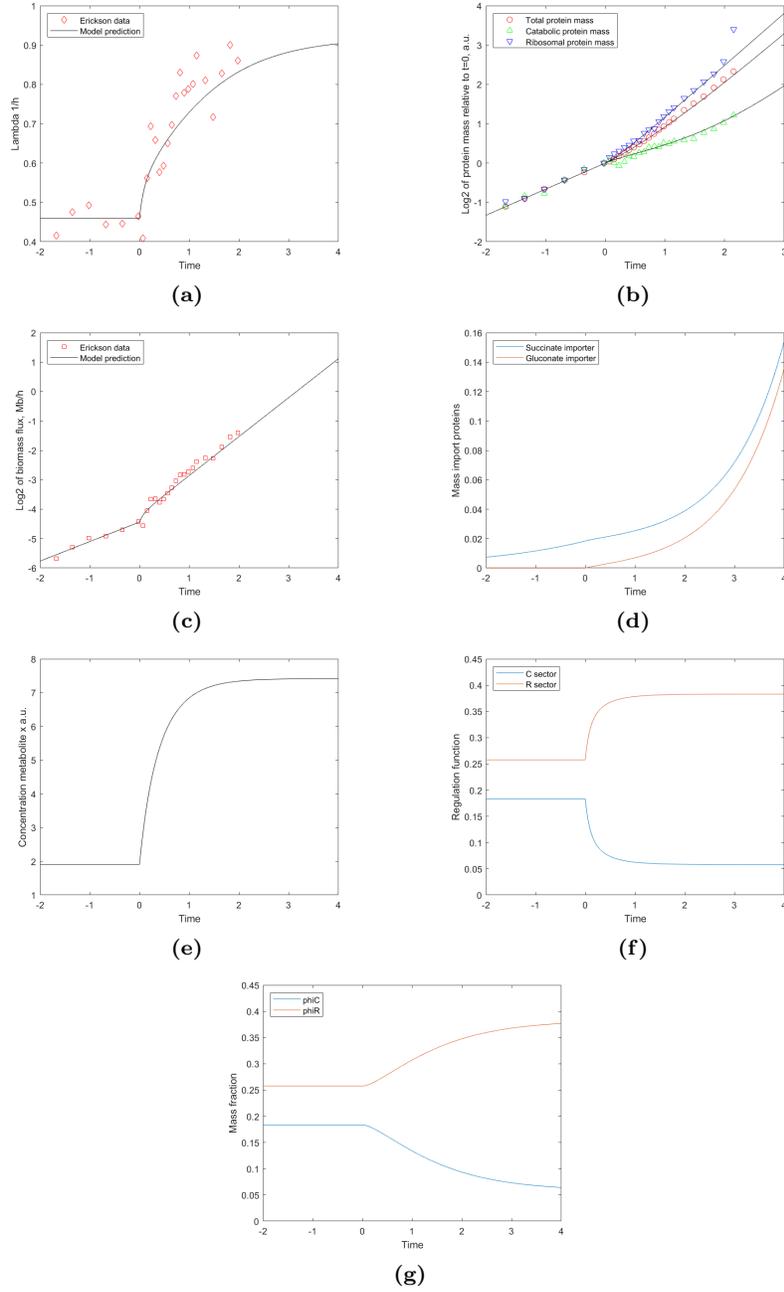
Figure 16: The ErTo model matches well to the gluconate upshift. Data from [11]. The initial substrate has succinate as the sole carbon source. At t = 0 gluconate is added. Parameters used: $\phi_{C,in} = 0.183301$, $\kappa_1 = 4.86084$, $\kappa_2 = 68.6763$. Lines are model predictions and dots are experimental measurements. Time is in hours. **a** Growth rate. **b** Protein masses. **c** Biomass flux. **d** Value $\tilde{M}_{C,i}$. **e** Concentration metabolite $\hat{\tilde{x}}$. **f** Regulation functions. **g** Mass fractions.

As the gluconate transporter starts being produced upon the addition of gluconate at $t = 0$ (figure 16D), the value of $\hat{x}$ increases (figure 16E) due to a higher carbon flux (equation 72). As a result $\mathcal{X}_C(\hat{x})$ deceases and $\mathcal{X}_R(\hat{x})$ increases (figure 16F). This causes the mass fractions of the C and R sector to decrease and increase respectively (figure 16G).

Our reproduction of the glucose downshift is shown in figure 17A, B and C. Our model clearly performs less well in matching the data from the glucose downshift. As figure 17A shows our model does not accuately reproduce the dynamics of the growth rate. The lowest value of the growth rate after the downshift is much higher in our model compared to the experimental data. The growth rate also recovers quite rapidly towards the post downshift steady-state. Because we overestimate the post shift growth rate we also overestimate the protein masses (figure 17B).
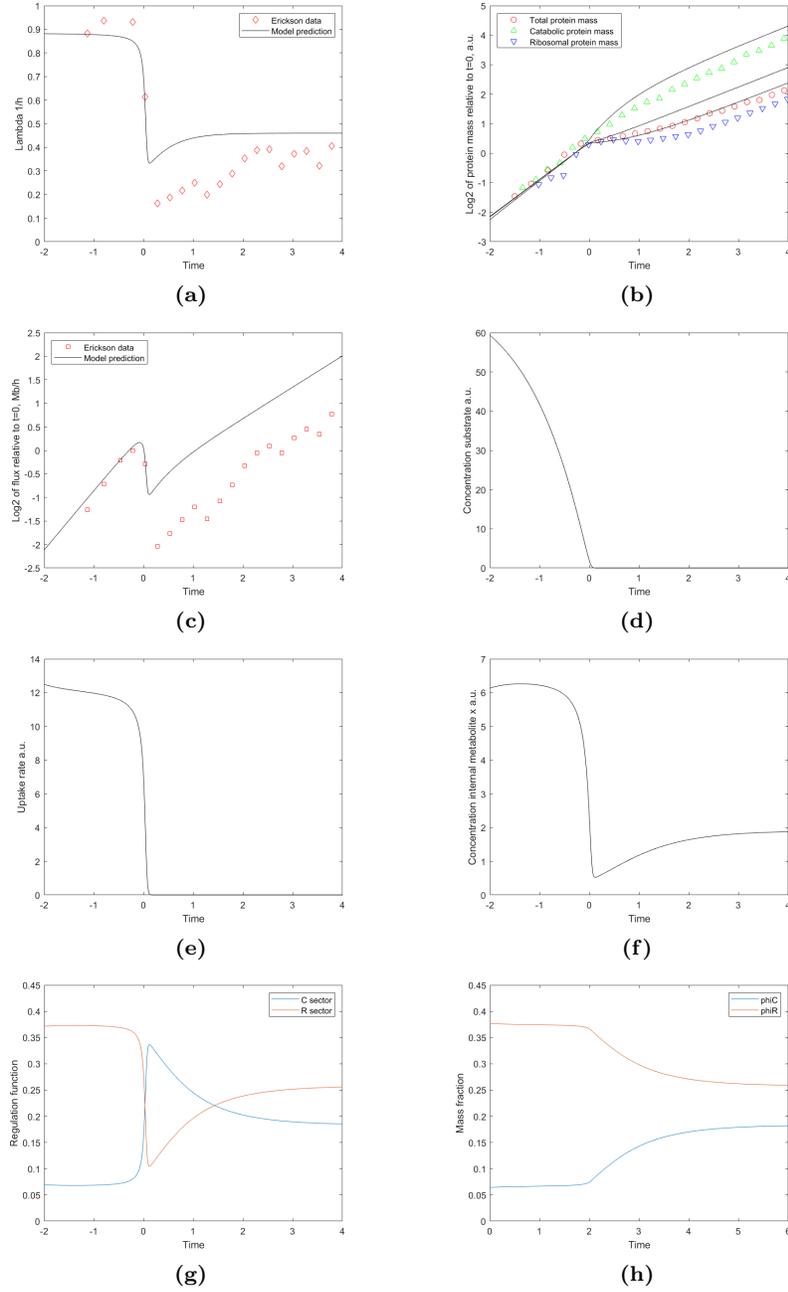
Figure 17: The ErTo model cannot match to the glucose downshift. Data from [11]. The initial substrate has both succinate and glucose as carbon source. Around t = 0 glucose is depleted. Parameters used: $\phi_{C,in} = 0.0643$, $\kappa_1 = 4.86084$, $\kappa_2 = 54.3813$, $K_{Ms,2} = 5$, $N_{2,in} = 475$, $V_E = 8$. Lines are model predictions and dots are experimental measurements. Time is in hours. **a** Growth rate. **b** Protein masses. **c** Biomass flux. **d** Concentration glucose. **e** Effective uptake parameter glucose importer. **f** Concentration metabolite $\hat{\bar{x}}$. **g** Regulation functions. **h** Mass fractions.

Around $t = 0$ glucose is depleted from the medium (figure 17D). As a result the uptake of glucose quickly decreases (figure 17E). This has the effect of decreasing $\hat{x}$ (figure 17f). The decreased value of $\hat{\hat{x}}$ causes $\mathcal{X}_C$ to increase and $\mathcal{X}_R$ to decrease (figure 17G). As a result ϕ_C and ϕ_R approach their new steady-state value. Our framework provides an intuitive explanation for the fact that the growth rate dips below its eventual steady-state value during downshifts. Because the uptake of glucose drops rapidly compared to the dynamics of the mass fractions there is a mismatch between ϕ_C and ϕ_R shortly after the downshift. As a result the R sector proteins drain the pool of the metabolite leading to a low growth rate (equations 44 and 48). As the mass fractions approach their optimal values $\hat{\hat{x}}$ recovers allowing for a faster growth rate.

Unlike the model of Erickson *et al*, the ErTo model is not capable of matching the dip in the growth rate by choosing a sufficiently low value for the initial condition of the V_E parameter (figure 9B).
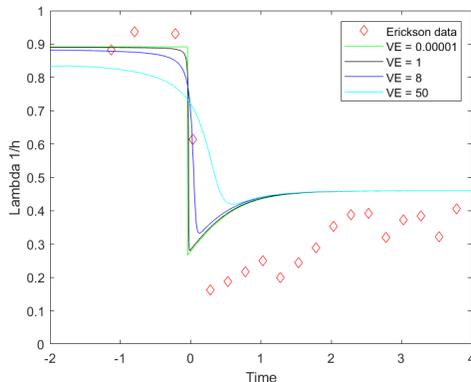
(a)

Figure 18: The ErTo model cannot provide a good match to the data of the glucose downshift, regardless of the value of V_E. Initial growth is on succinate and glucose. Around t = 0 glucose is depleted. Black lines are model prediction, red dots are experimental data. Other parameters are: $\phi_{C,in} = 0.0643$, $\kappa_1 = 4.86084$, $\kappa_2 = 54.3813$, $K_{Ms,2} = 5$, $N_{2,in} = 475$.

The ErTo model has another issue that is not directly obvious from figures 16 and 17, but becomes apparent when modelling up- or downshifts in which $\bar{\lambda}_{1,2} \gg max(\bar{\lambda}_1, \bar{\lambda}_2)$. For the upshift in which glycerol is added to a medium already containing succinate the experimentally measured steady-state growth rates are $\bar{\lambda}_1 = 0.46$, $\bar{\lambda}_2 = 0.62$ and $\bar{\lambda}_{1,2} = 0.78$. For this system the ErTo model severely underestimates $\bar{\lambda}_{1,2}$ (figure 19). The ErTo model also provides an unsatisfactory match to the data of the glycerol downshift (figure 20).
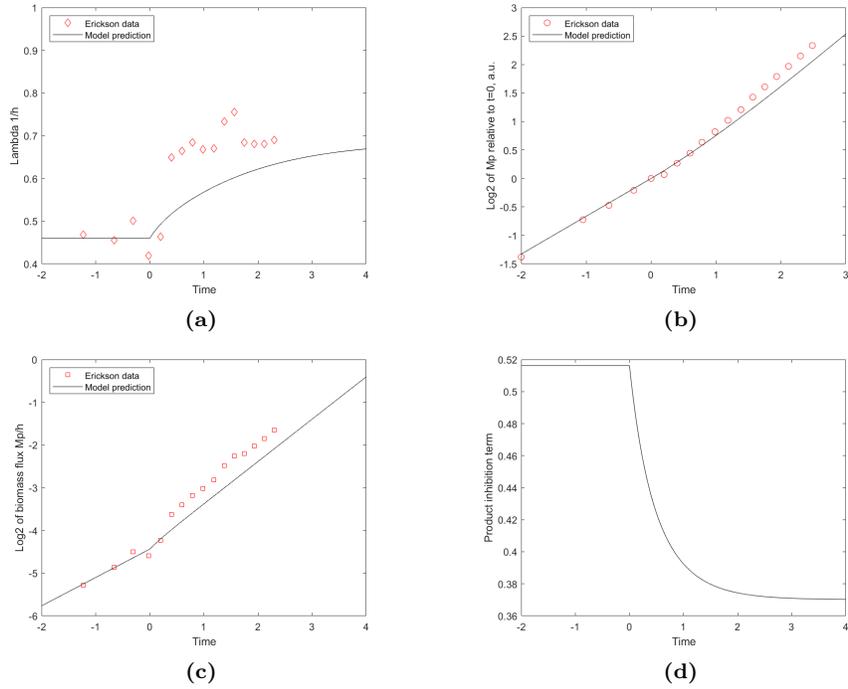
Figure 19: ErTo model cannot fit the glycerol upshift. Data from Erickson *et al.* The initial substrate has succinate as the sole carbon source. At t = 0 glycerol is added. Lines are model predictions, dots are experimental measurements. Parameters used: $\phi_{C,in} = 0.183301$, $\kappa_1 = 4.86084$, $\kappa_2 = 11.1712$. **a** Growth rate. **b** Total protein mass. **c** Biomass flux. **d** Product inhibition term $\frac{\hat{K}_I}{\hat{K}_I + \hat{x}}$.
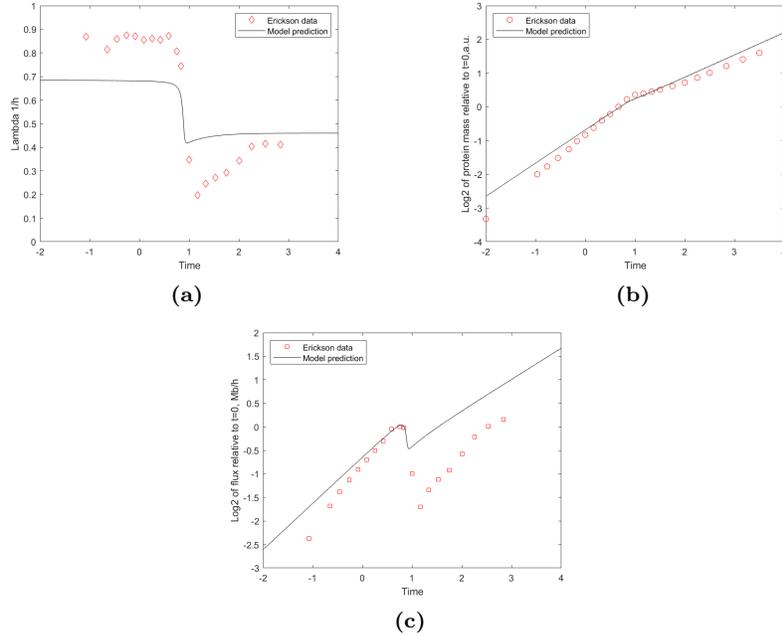
Figure 20: ErTo model cannot fit the glycerol downshift. Data from Erickson *et al*. The initial substrate has succinate as a carbon source. Around t = 0 glycerol is depleted from the medium. Lines are model predictions, dots are experimental measurements. Other parameters are $\phi_{C,in} = 0.1157$, $\kappa_1 = 4.86084$, $\kappa_2 = 11.1712$, $K_{Ms,2} = 5.6$, $N_{2,in} = 250$, $V_E = 1$. **a** Growth rate. **b** Protein mass. **c** Biomass flux.

In the supplemental materials we show that $\bar{\lambda}_{1,2}$ is higher in the glycerol downshift when compared to $\bar{\lambda}_{1,2}$ in the upshift experiment or the $\bar{\lambda}_{1,2}$ predicted by the composition formula. We show there that by modelling glycerol as if it was a medium allowing for a higher steady-state growth rate, and by adjusting the initial condition $N_{2,in}$, the Erickson *et al* model was able to match the data from the experiment. This is not the case for the ErTo model (figure 21A), even with an extremely abrupt (low V_E) downshift (figure 21B).

50
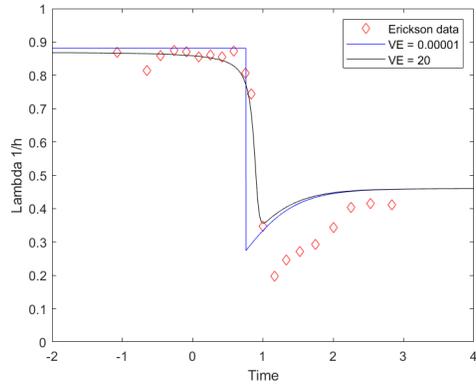
(a))

Figure 21: The predicted minimum of the growth rate is not low enough for any value of V_E to match the data of the glycerol downshift. Lines are model predictions. Dots are experimental measurements. The initial substrate has succinate as a carbon source. Around t $= 0$ a hypothetical substrate is depleted from the medium. Other parameters are $\phi_{C,in} = 0.0666$, $\kappa_1 = 4.86084$, $\kappa_2 = 50$, $K_{Ms,2} = 5.6$, $N_{2,in} = 1000$.

5 Discussion

We have derived a model of optimal resource allocation between the C- and R sector in *E. coli*. Our model can in principle describe both steady state behaviour and growth transitions. In steady state our model is quite similar to the model by Towbin *et al.* After all, we use similar equations for $\frac{\mathrm{d}\phi_C}{\mathrm{d}t}$ and $\frac{\mathrm{d}x}{\mathrm{d}t}$ and the form of the regulation functions we derived through optimality is very close to the regulation functions postulated by Towbin *et al*. Towbin *et al* did not derive their regulation functions but instead specified a general form *a priori*. They then fit this general form to the c-line. From this general form Towbin *et al* derived conditions under which regulation is optimal and showed experimentally that these conditions generally hold. However, their analysis rests completely on the functional form of the regulation functions. If this is far from optimal even under the best fit to the c-line, the optimal conditions derived by Towbin *et al* represent nothing more than the optimal way to be sub optimal. From the fact that the regulation function of Towbin *et al* are indeed close to our optimal regulation functions we conclude that the optimality conditions derived by Towbin *et al* could potentially be generalised beyond the chosen form of their regulation functions.

In contrast to the Towbin *et al* model the ErTo model is capable of describing transitions between steady states. We applied it to the upshift and downshift experiments performed by Erickson *et al*. We found that the ErTo model performs less well than the model by Erickson *et al*. The ErTo model is not capable of correctly predicting the steady state growth rate on the two-substrate media. We suspect this could be because we assume that product inhibition limits the flux through the import proteins. We calculate the value of the parameter $\hat{\kappa}_i$ (which represents the quality of the environment) using equation 68. This equation gives the value of $\hat{\kappa}_i$ which results in the correct steady state growth rate, taking into account the value of the product inhibition term $\frac{\hat{K}_I}{\hat{K}_I + \hat{x}}$. However, for e.g. the succinate glycerol downshift the value of the product inhibition term is significantly lower on the two-substrate medium than on either of the single substrates (figure 19)D. The values of κ_1 and κ_2 are not "adapted" to the increased level of product inhibition, resulting in a lower prediction of the two-substrate growth rate when compared to the prediction of the composition formula (equation 40). The model by Erickson *et al* does not have this problem, since it does not model product inhibition. We cannot resolve this by simply excluding product inhibition from our model. We redid the fitting procedure described in section 4.1.6 with the additional constraint that $\hat{K}_I \geq 100$ (In this regime there is little product inhibition). In consequence, however, the ErTo model is now unable to provide a good fit to the Towbin *et al* data (figures 22 and 23). Particularly the fit to the O-curve data declines in quality. Interestingly, the c-line becomes very close to linear for the entire range of growth rates under this regime. Whether the fit to the downshift experiments increases in

quality for these parameters is a promising direction for future research that can be easily implemented.
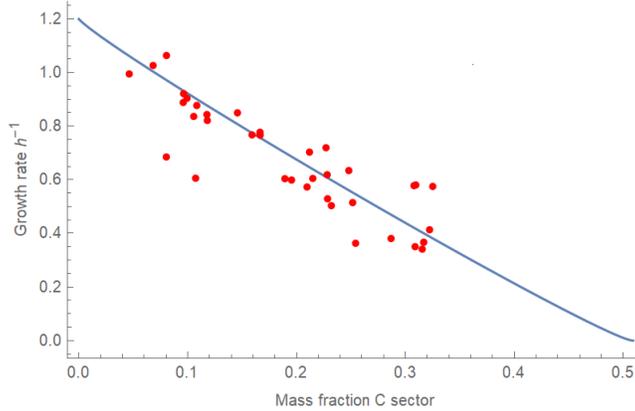

Figure 22: C-line becomes as good as linear when ErTo model is fit with the additional constraint that $\hat{K}_I \geq 100$. Fitted parameters: $\phi_Q = 0.49111$, $\hat{K}_I = 100$, $\hat{\gamma} = 2.35808$, $\theta = 0.520607$.
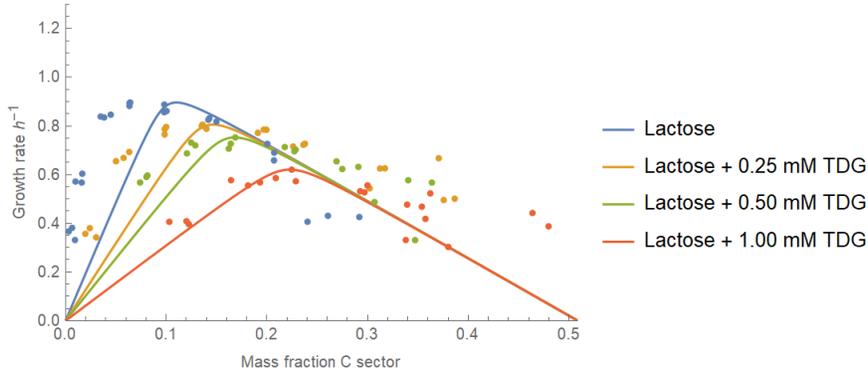

Figure 23: The ErTo model can no longer fit to the O-curve data from Towbin *et al* under the constraint that $\hat{K}_I \geq 100$. *E. coli* grown on lactose with with varying amounts of the lactose transport inhibitor TDG added to the medium. Fitted parameters: $\phi_Q = 0.49111$, $\hat{K}_I = 100$, $\hat{\gamma} = 2.35808$, $\theta = 0.520607$.

We have assumed that the value of $\hat{K}_I$ is constant between substrates. This is not necessarily the case in the real biological system. Indeed, Towbin *et al* chose to fit this parameter separately for each substrate. This would naturally lead to the correct steady state growth rate on each medium. However, because our regulation functions depend on the value of $\hat{K}_I$, we cannot fit this parameter separately for each environment while maintaining optimality.

A possible solution to the problem that the ErTo model cannot match the two-substrate growth rate is to explore whether an intermediate value of $\hat{K}_I$ allows the ErTo model to maintain a reasonable quality of fit to the O-curve data, while simultaneously lessening the effect of product inhibition to such an extent that the ErTo model can describe the upshift and downshift experiments more successfully. This has not yet been done, but is a potential future direction of research.

Another interesting direction of research is to investigate whether the ErTo model can more accurately model growth transitions if it uses the regulation functions of Erickson *et al* (equations 27 and 26). Because these are derived from the growth laws the steady state growth rate on the two-substrate media should follow from the composition formula. We would need to investigate the nature of the translational activity within the ErTo model framework to do this.

Even if the ErTo model were to match the growth rate on the two-substrate medium it would still be unable to match the data from the downshift experiments. In the prediction of the ErTo model, the growth rate does not dip low enough after the shift and the recovery of the growth rate towards the new steady state is too quick compared to experimental data. Apart from the assumption of product inhibition and michaelis menten kinetics by the intermediate metabolite we also constrain the sum of ϕ_C and ϕ_R to be constant (equation 42). By using the regulation functions of Erickson *et al*, we would also be letting go of this assumption. *E. coli* has many more protein sectors than just the C and R sector. The sum constraint (equation 42) may be too inflexible to allow accurate modelling of the real biological system. The regulation functions of the ErTo model cannot be derived without the sum constraint assumption.

While the failures of the ErTo model on the dynamical data could be the consequence of any of the issues above, it could also be the result of the fact that the Towbin *et al* and Erickson *et al* data are fundamentally incompatible. After all, the experiments were performed in different labs using different methods. Towbin *et al* grew *E. coli* in plate wells, whereas Erickson *et al* used batch culture. It is possible that trying to match the dynamical data of Erickson *et al* using a model fit to the data from Towbin *et al* constitutes a hopeless endeavour. In this case measurements from a single lab for the c-line data, O-curve data and dynamical data would be necessary.

The ErTo model is still very much a work in progress. It has already allowed us to vindicate the approach taken by Towbin *et al*, showing that their model can indeed make general conclusions regarding optimality. Resolving the problem that the ErTo model cannot accurately predict the two-substrate growth rate could potentially lead to increased understanding of the link between product inhibition and regulation in *E. coli*. Investigating why Erickson *et al* appear to match the dynamics of downshifts better, even if the two-substrate growth rate

is correct, could lead to further biological insights.

References

[1] Karl Kochanowski et al. "Few regulatory metabolites coordinate expression of central metabolic genes in Escherichia coli". In: *Molecular Systems Biology* 13.1 (2017), p. 903.

[2] Sheng Hui et al. "Quantitative proteomic analysis reveals a simple strategy of global resource allocation in bacteria". In: *Molecular systems biology* 11.2 (2015), p. 784.

[3] Conghui You et al. "Coordination of bacterial proteome with metabolism by cyclic AMP signalling". In: *Nature* 500.7462 (2013), pp. 301–306.

[4] Matteo Mori et al. "From coarse to fine: the absolute Escherichia coli proteome under diverse growth conditions". In: *Molecular systems biology* 17.5 (2021), e9536.

[5] Matthew Scott et al. "Emergence of robust growth laws from optimal regulation of ribosome synthesis". In: *Molecular systems biology* 10.8 (2014), p. 747.

[6] Alexander Schmidt et al. "The quantitative and condition-dependent Escherichia coli proteome". In: *Nature biotechnology* 34.1 (2016), pp. 104–110.

[7] Tomohiro Shimada et al. "Novel roles of cAMP receptor protein (CRP) in regulation of transport and metabolism of carbon sources". In: *PloS one* 6.6 (2011), e20081.

[8] Justin J Lemke et al. "Direct regulation of Escherichia coli ribosomal protein promoters by the transcription factors ppGpp and DksA". In: *Proceedings of the National Academy of Sciences* 108.14 (2011), pp. 5712–5717.

[9] Matthew Scott et al. "Interdependence of cell growth and gene expression: origins and consequences". In: *Science* 330.6007 (2010), pp. 1099–1102.

[10] Benjamin D Towbin et al. "Optimality and sub-optimality in a bacterial growth law". In: *Nature communications* 8.1 (2017), pp. 1–8.

[11] David W Erickson et al. "A global resource allocation strategy governs growth transition kinetics of Escherichia coli". In: *Nature* 551.7678 (2017), pp. 119–123.

[12] Martijn Wehrens. "Dynamical regulation in single cells". PhD thesis. Delft University of Technology, 2019.

[13] Markus Basan et al. "Inflating bacterial cells by increased protein synthesis". In: *Molecular systems biology* 11.10 (2015), p. 836.

[14] Karl Kochanowski et al. "Global coordination of metabolic pathways in Escherichia coli by active and passive regulation". In: *Molecular systems biology* 17.4 (2021), e10064.

[15] Robert Planqué et al. "Maintaining maximal metabolic flux by gene expression control". In: *PLOS Computational Biology* 14.9 (2018), e1006412.

[16] Rutger Hermsen et al. "A growth-rate composition formula for the growth of E. coli on co-utilized carbon substrates". In: *Molecular systems biology* 11.4 (2015), p. 801.

6 Supplementary Materials

Applying the Erickson *et al* model to the succinate/glycerol up- and downshift

To show that the Erickson *et al* model works on more than one up- or downshift, we reproduce the match of the Erickson *et al* model to the data of the glycerol upshift (figure 24) and glycerol downshift (figure 25) experiments. While the model predicts the glycerol upshift well, this is not true for the downshift experiment. This appears largely due to the fact that the growth rate on the co-utilized substrate in the glycerol downshift is much higher than the growth rate predicted by the composition formula or measured experimentally by Erickson *et al* in the glycerol upshift. The medium used for the glycerol downshift may have been contaminated allowing for a higher steady-state growth rate. To test this hypothesis we model the experiment as a downshift from succinate and a second hypothetical carbon substrate. This hypothetical substrate sustains a steady-state growth rate of $\bar{\lambda} = 0.82$ under carbon limiting conditions. From the composition formula it follows that on the co-utilized substrate $\bar{\lambda}_{1,2} = 0.876822$, which matches the co-utilized growth rate observed in the glycerol downshift experiment. Because of the higher co-utilized growth rate glycerol is depleted at a faster rate. As a result the value of $N_{2,in}$ needs to be set at a higher value to ensure correct timing of the downshift. Using only the new values of λ_2 and $N_{2,in}$ the dynamics which follow from the model once again provide a good match to the experimental data. Presumably Erickson *et al* did the same, because they show a good model fit to the glycerol downshift.

(a)

(b)

(c)
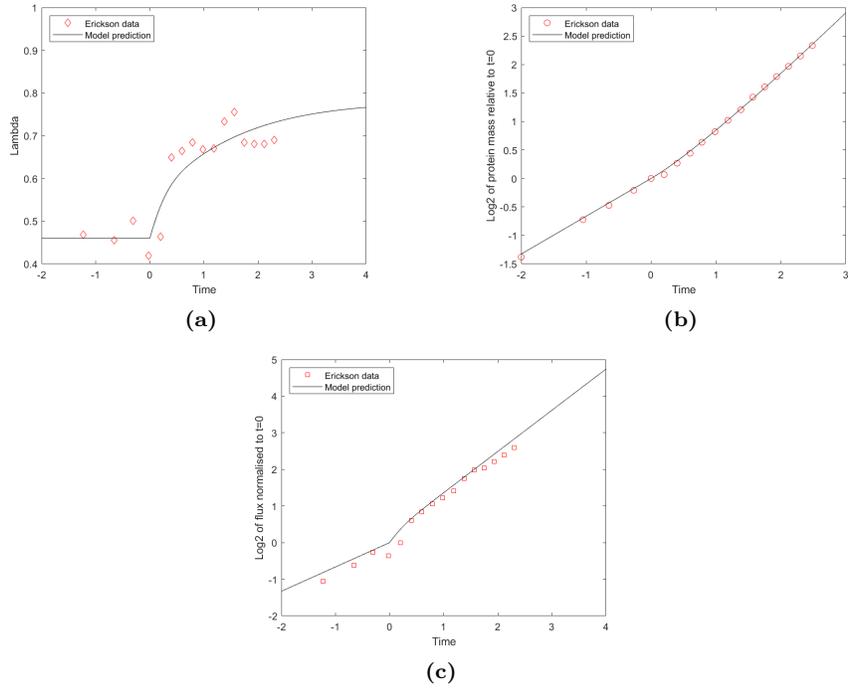
Figure 24: The Erickson *et al* model provides a good match to the update from the glycerol upshift experiment. Initial growth is on succinate. At t = 0 glycerol is added. Data from paper [11]. Line is model prediction and dots are experimental measurements. Parameters used: $\bar{\lambda}_1 = 0.46$, $\bar{\lambda}_2 = 0.62$. Time is in hours. **a** Growth rate. **b** Total protein mass. **c** Biomass flux.
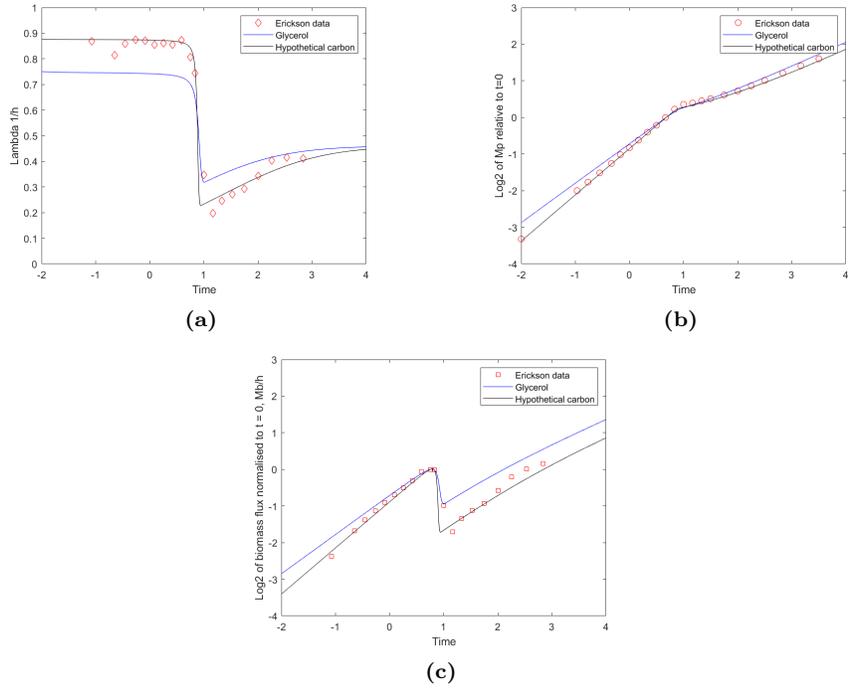
Figure 25: The Erickson et al model can match the glycerol downshift by assuming growth is on a hypothetical carbon substrate with $\bar{\lambda}_2 = 0.82$. Initial growth is on succinate and glycerol/hypothetical substrate. Around $t = 0$ glycerol/hypothetical substrate is depleted. Data from paper [11]. Parameters used: $\bar{\lambda}_1 = 0.46$, $\bar{\lambda}_2 = 0.62$ (succinate) $\bar{\lambda}_2 = 0.82$ (hypothetical substrate), $N_{2,in} = 157$ (succinate) $N_{2,in} = 269$ (hypothetical substrate), $V_E = 0.6$ (for both). Time is in hours. **a** Growth rate. **b** Total protein mass. **c** Biomass flux.