Modeling and analysis of a gene-regulatory feed-forward loop with basal expression of the second regulator

Louisa Roselius[a], Dirk Langemann[b], Johannes Müller[c,d], Burkhard A. Hense[d], Stefan Filges[a], Dieter Jahn[a], Richard Münch[a,*],

[a]*Institute of Microbiology and Braunschweig Integrated Centre of Systems Biology, Technische Universität Braunschweig, Spielmannstr. 7, D-38106 Braunschweig, Germany*
[b]*Institute of Computational Mathematics, Pockelsstr. 14, Technische Universität Braunschweig, D-38106 Braunschweig, Germany*
[c]*Centre for Mathematical Sciences, Technische Universitat München, Boltzmannstr. 3, D-85747 Garching/Munich, Germany*
[d]*Institute of Computational Biology, Helmholtz Zentrum München - German Research Center for Environmental Health, Ingolstädter Landstr. 1, D-85764 Neuherberg, Germany*

Abstract

Efficient adaptation strategies to changing environmental conditions are essential for bacteria to survive and grow. Fundamental restructuring of their metabolism is usually mediated by corresponding gene regulation. Here, often several different environmental stimuli have to be integrated into a reasonable, energy-efficient response. Fast fluctuations and overshooting have to be filtered out. The gene regulatory networks for the anaerobic adaptation of the pathogenic bacterium *Pseudomonas aeruginosa* is organized as a feed-forward loop (FFL), which is a three-gene network motif composed of two transcription factors (Anr for oxygen, NarxL for nitrate) and one target (Nar for nitrate reductase). The upstream transcription factor (Anr) induces the downstream transcription factor (NarXL). Both regulators act together positively by inducing the target (Nar) via a direct and indirect regulation path (coherent type-1 FFL). Since full promoter activity is only achieved when both transcription factors are present the target operon is expressed with a delay. Thus, in response to environmental stimuli (oxygen, nitrate), signals are mediated and processed in a way that short pulses are filtered out. In this study we analyze a special kind of FFL called FFL_k by means of a family of ordinary differential equation models. The secondary FFL regulator (NarXL) is expressed constitutively but further induced in presence of the upstream stimuli. This FFL modification has substantial influence on the response time and cost-benefit ratio mediated by environmental fluctuations. In order to find conditions where this regulatory network motif might be beneficial, we analyzed various models and environments. We describe the observed evolutionary advantage of FFL_k and its role in environmental adaptation and pathogenicity.

Keywords: gene regulation, feed-forward loop, mathematical modeling, environmental adaptation, energy metabolism

1. Introduction

Most bacteria are regularly exposed to changing physical and chemical conditions in their environment. In order to survive in a wide range of habitats, it is necessary to adapt to these environmental fluctuations, which is often mediated by transcription regulation networks. In response to external signals, transcription factors are able to change the transcription rate of target genes, which encode proteins or RNAs required for efficient adaptation. Usually multiple signals and regulators are integrated at the target gene promotor in a complex gene regulatory network (GRN). Global analysis of the design principles of microbial GRNs revealed overrepresented sub-networks, so called network motifs (Shen-Orr et al., 2002; Alon, 2007).

One of the most abundant network motifs found in microbial systems is the feed-forward loop (FFL). Examples for FFLs are used for the regulation of various sugar utilization systems of *Escherichia coli* (Mangan et al., 2003, 2006; Görke and Stülke,

2008), of flagella production (Kalir et al., 2005) and the induction of the *m*-xylene biodegradation pathway in *Pseudomonas putida* (Silva-Rocha and de Lorenzo, 2011). A nested system of feed-forward loops is found in the sporulation pathway of *Bacillus subtilis* (Eichenberger et al., 2004). Structurally, a FFL is a three-gene network of two transcription factors x and y that act together to regulate the target gene(s)/protein(s) z with the additional property that y is also regulated by x. Thus, there is a direct and indirect path to regulate z. Both regulators usually respond to the external stimuli s_x and s_y, respectively (Fig. 1A). Since there are three types of regulatory interactions, which can be either positive or negative (activation or repression), a FFL motif offers eight different possibilities of structural configurations (Mangan and Alon, 2003). The relationship between each FFL topology with its dynamic response behaviour shows a unique probability distribution pattern with differences in its intrinsic plasticity and robustness (Macia et al., 2009; Widder et al., 2012).

In our study, we are interested in the so-called type-1 coherent FFL where all regulatory interactions are positive. It was shown that this type of FFL has the potential to act as a delay element since the second regulator has first to accumu-
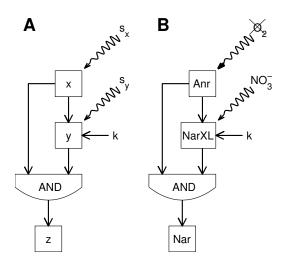
Figure 1: A) General structure of a type-1 coherent feed-forward loop motif. Regulator x activated by stimulus s_x induces the expression of regulator y which is activated by stimulus s_y. Both regulators act together in inducing the target gene(s) and corresponding protein(s) designated as z. Optionally, y has a basal expression rate of k. B) Anaerobic induction of the denitrification genes of *P. aeruginosa* in presence of nitrate and absence of oxygen via the regulators Anr and the two-component system NarXL. NarXL has a basal half-maximal expression which is presented by the parameter k.

late for successful transcription of the target gene(s) (Mangan and Alon, 2003; Kalir et al., 2005; Wall et al., 2005). Consequently, these circuits can act as filter for unwanted responses to fluctuating inputs. In contrast to previous studies we analyze a special case of the type-1 coherent FFL where the second regulator y shows a significant basal expression level denoted by k. Therefore, we named this regulatory circuit FFL_k. While FFL_k leads to a permanent cost factor due to the production of the regulator y it is expected that the pulse filtering is reduced to shorter pulses since the delay time to the onset of z is significantly decreased. FFL_k-type regulation was previously described for *E. coli*, where the arabinose responsive regulator AraC shows a significant basal expression level (Mangan et al., 2003). In this study we refer to an FFL_k involved in anaerobic adaptation of the opportunistic pathogen *P. aeruginosa*. During an environmental shift to anoxic conditions this organism can efficiently adapt to dissimilatory denitrification which is a respiratory process where nitrate is used instead of oxygen as electron acceptor for energy generation. While, oxygen is the energetically most favourable electron acceptor, dissimilatory nitrate respiration is the system with highest energy yields in anaerobic environments (Strohm et al., 2007). The two environmental signals "loss of oxygen" and "presence of nitrate" are sensed by the regulator Anr and the two-component system NarXL, respectively. It has been shown in previous analyses, that the *narXL* operon is constitutively expressed under aerobic conditions and is induced about two-fold after anaerobic induction (Fig. 1B) (Schreiber et al., 2007). We argue, that the resulting decreased response time to anaerobic pulses might be a tradeoff between the essential need for energy production versus the cost-intensive biosynthesis of the nitrate reducing enzyme apparatus. Such a balancing of cost and increased fitness

seems to play a major role in the evolutionary selection of protein expression levels and network motifs (Lan and Tu, 2013; Dekel et al., 2005; Shachrai et al., 2010).

In this context, we analyzed various models in order to find conditions where a FFL_k is generally beneficial in comparison to a FFL without a basal expression. An optimal net-benefit relation with a non-trivial, i. e. non-vanishing and non-complete, basal expression mirrors an evolutionary advantage of a FFL_k. The paper is organized as follows: First the modeling approach of the feed-forward loop is generally introduced in Sec. 2. In the following Sec. 3 a linear modeling approach is discussed. There, we show that the use of linear influence functions is not suitable to demonstrate a non-trivial optimal net-benefit relation. The first argument for an evolutionary advantage is found in Sec. 4 using a nonlinear model with an optimal net-benefit function. However, the short delay time and ability of fast target production accounts for significant costs in this model. Directly after a signal switch to anaerobic phases the instantaneous net-benefit of this model is negative, which is not suitable to explain an evolutionary advantage because such a situation would eventually lead to cell death. This model is able to explain the advantage of the basal expression but not the evolutionary stability of the whole system. Consequently, Sec. 5 extends the model by an energy storing mechanism, that integrates the mentioned costs and turns their utilization into the advantage of a rapid change to anaerobic respiration. This model with an energy store allows for a better understanding of the essential role of the basal expression. However, the use of linear influence functions again does not provide arguments for any optimality of a non-vanishing and non-complete basal expression. Finally, the model with the energy store and the nonlinearities were combined, and we will present a simple mathematical model underlining the evolutionary advantage of the basal expression. We conclude with numerical simulation results and a discussion in the context of anaerobic adaptation and pathogenicity of *P. aeruginosa*.

2. Modeling the feed-forward loop

The generalized feed-forward-loop model contains a system of three differential equations corresponding to the two regulators x and y and the target z. For simplicity, transcription and translation were modeled as one process, so the state variables represent protein levels. All state variables are normalized to one, $\max x = \max y = \max z = 1$, where the value 1 means maximum abundance of the respective protein. The degradation rates of the regulators x and y and the target z are determined by their own concentrations and assumed to be identical. A normalization of the time scale gives the dynamical system

$$\dot{x} = s_x - x, \tag{1}$$
$$\dot{y} = g_k(x) - y, \tag{2}$$
$$\dot{z} = m(x, y) - z. \tag{3}$$

The first regulator x is induced by the signal s_x. The signal s_y is assumed to be constant. Here, $s_x = s_x(t)$ is simplified as a

piecewise constant function, which switches between the values 0 for no stimulus and 1 for activating stimulus. All points of switching between zero and one are called T_i with $T_{i+1} > T_i$ and $T_0 = 0$.

$$s_x(t) = \begin{cases} 0 & \text{if } t \in [T_{2i}, T_{2i+1}[, \\ 1 & \text{if } t \in [T_{2i+1}, T_{2i+2}[, \ i \in \mathbb{Z}. \end{cases} \quad (4)$$

In the following T_1 is the time of the first switch from no stimulus to an activating stimulus. The production of the second regulator y depends on the regulation $g_k(x)$, which includes the level of basal expression $k \in [0, 1]$ and the first regulator x. Finally, the target z is induced by x and y via the function $m(x, y)$.

The regulations $g_k(x)$ and $m(x, y)$ are monotonously increasing functions in each component. Since a full presence of the regulators generate a maximized expression, we get the conditions

$$g_k(0) = k, \ g_k(1) = 1 \ \forall k \in [0, 1] \ \text{and} \ m(1, 1) = 1. \quad (5)$$

A higher basal expression provokes a higher regulation of the transcription factor y, i. e. $g_k(x) > g_\ell(x) \geq 0$ is true for $k > \ell$ and for all $x \in [0, 1]$. So, g_k needs to be monotonously increasing in k.

The net-benefit C is produced by the benefit minus the costs of the feed-forward-loop. The production and maintenance of regulator y and the target z is an energy consuming process with cost factors η_y and η_z, respectively. However, the availability of z is beneficial and increases the fitness under the appropriate environmental condition when the signal s_x is present.

Therefore, we define the net-benefit by

$$C = \lim_{T \to \infty} \frac{1}{T} \int_0^T s_x z - \eta_y g_k(x) - \eta_z m(x, y) \, \mathrm{d}t. \quad (6)$$

The net-benefit is well defined whenever the limit on the right hand side exists. We show below, that periodic excitations s_x lead to periodic x, y, z, and thus, T-periodic solutions of system (1 - 3) provide the net-benefit functional

$$C = \frac{1}{T} \int_0^T s_x z - \eta_y g_k(x) - \eta_z m(x, y) \, \mathrm{d}t. \quad (7)$$

The integrand $c_k = s_x z - \tilde{c}_k = s_x z - \eta_y g_k(x) - \eta_z m(x, y)$ in Eq. (7), i. e. the time-dependent difference of cost and benefit, is called the instantaneous net-benefit with $\tilde{c}_k$ as instantaneous cost.

In order to discuss the influence of the basal expression k to the net-benefit and a possible optimal k, we need some preliminary facts about the system (1 - 3) formulated in the following lemmas.

Lemma 2.1.
The differential equation $\dot{u} = f - u$ with the excitation f has the solution

$$u = \mathrm{e}^{-t} u(0) + \int_0^t \mathrm{e}^{\theta - t} f(\theta) \, \mathrm{d}\theta. \quad (8)$$

A T-periodic excitation leads to the periodic steady state

$$L[f] = u = \frac{\mathrm{e}^{-t}}{1 - \mathrm{e}^{-T}} \int_0^T \mathrm{e}^{\theta - T} f(\theta) \, \mathrm{d}\theta + \int_0^t \mathrm{e}^{\theta - t} f(\theta) \, \mathrm{d}\theta. \quad (9)$$

Proof. A simple calculation proves the lemma. □

The periodic steady state $L[f] = u$ in Eq. (9) defines the linear operator $L : \mathcal{L}_1([0, T]) \to \mathcal{L}_1([0, T])$.

We emphasize the dependence of the solution on the basal expression by an index k. The periodic case provides $x = L[s_x]$, $y_k = L[g_k(x)]$ and $z_k = L[m(x, y_k)]$.

Lemma 2.2. *Solutions (x, y_k, z_k) and (x, y_ℓ, z_ℓ) of system (1 - 3) with $k \geq \ell \geq 0$ and $y_k(0) \geq y_\ell(0)$ and $z_k(0) \geq z_\ell(0)$ obey $y_k(t) \geq y_\ell(t)$ and $z_k(t) \geq z_\ell(t)$ for all $t > 0$.*

Proof. Due to Eq. (8) in Lemma 2.1, the monotonicity of $g_k(x)$ in k implies the monotonicity of the solution $y_k(t)$ in k. Furthermore, $y_k(t)$ depends monotonously increasing on the initial value $y_k(0)$.

Analogously, the monotonicity of the excitation $m(x, y_k)$ in y_k leads to a target $z_k(t)$ which increases monotonously with $z_k(0)$ and with k. □

In the case of a periodic signal $s_x(t)$, any transient phase fades out, and the analogous argumentation shows the monotonicity of $y_k = L[g_k(x)]$ and $z_k = L[m(x, y_k)]$ in k. Therefore, an organism with basal expression, i. e. with $k > 0$ has a higher concentration of the regulators y_k and thus a higher expression of the target z_k for every periodic steady state and for every system behavior with identical initial conditions.

3. Linear model

The simplest cases for $g_k(x)$ and $m(x, y)$ are linear functions in each component, so we call the model (1 - 3) the linear model by using

$$g_k(x) = k + (1 - k)x \ \text{and} \ m(x, y) = xy. \quad (10)$$

Lemma 3.1. *The first switch at T_1 and the time after the switch τ in the system (1 - 3) with the initial values $x(0) = 0$, $y_k(0) = k$ and $z_k(0) = 0$ induces the growth behavior*

$$x(T_1 + \tau) = O(\tau) \ \text{and} \ y_k(T_1 + \tau) = k + O(\tau^2).$$

The additional assumption $m(x, y) = O(xy)$ for small x and y assures

$$z_k(T_1 + \tau) = kO(\tau^2) + O(\tau^4).$$

Proof. A calculation using the linear constitutive functions in Eq. 10 and the Hartman-Grobman theorem provide the proof. □

Lemma 3.1 shows that in all situations the instantaneous costs $\tilde{c}_k = \eta_y g_k(x) + \eta_z m(x, y_k) = \eta_y k + O(\tau)$ are much higher for

small τ than the benefit $s_x(T_1 + \tau)z_k(T_1 + \tau)$ shortly after the first switch (Fig. 2 D).

The FFL model (Eqs. 1 - 3) results in a negative instantaneous net-benefit c_k at $t = T_1 + \tau$ for small $\tau > 0$, see Fig. 2D, and thus, it cannot explain the evolutionary stability of this bacterial system. Despite the negative instantaneous net-benefit, the model is tested for optimality in the basal expression k.

In the following, we prove that the net-benefit functional (7) cannot have a strong maximum with $k \in\ ]0, 1[$ if the linear activations (10) are used. The idea behind this is, to show that $C = C(k)$ is a linear function in k. For that purpose, the linear model cannot explain a selective advantage for FFL_k compared with FFL_0 under the given assumptions.

Theorem 3.1.
The linear model has no strongly optimal k in the net-benefit $C = C(k)$.

Proof.
Using the linearity of the operator L and Eq. (10), we write the solution of system (1 - 3) as

$$x = L[s_x] \text{ and } y_k = L[k + (1 - k)\, x].$$

With $y_0 = L(x)$ in the absence of a basal expression, we find

$$y_k = k + (1 - k)\, y_0 \text{ and}$$
$$z_k = L[xy_k] = L[x\,(k + (1 - k)\, y_0)]$$
$$= ky_0 + (1 - k)\, L[y_0 x].$$

The net-benefit (7) with these calculated solutions reads as

$$C(k) = \frac{1}{T} \int_0^T s_x\,(ky_0 + (1 - k)\, L[y_0 x]) - \dots$$
$$- \eta_y\,(k + (1 - k)\, x) - \eta_z x\,(k + (1 - k)\, y_0)\ \mathrm{d}t,$$

where $x = L[s_x]$ and $y_0 = L[x] = L[L[s_x]]$ are independent of k. Therefore, the net benefit has the form

$$C(k) = \frac{1}{T} \int_0^T Q_1(s_x, \eta_y, \eta_z) \cdot k + Q_2(s_x, \eta_y, \eta_z)\ \mathrm{d}t.$$

with the time-dependent coefficients

$$Q_1 = s_x(y_0 - L[y_0 x]) - \eta_y(1 - x) - \eta_z x(1 - y_0)$$

and

$$Q_2 = s_x L[y_0 x] - \eta_y - \eta_z x.$$

The net-benefit functional is a linear function in k. Hence, strong optima of the net-benefit occur only at zero or one, depending on the sign of Q_1. □

A typical simulation result for a fluctuating environment with cyclic changes between zero stimulus and activating stimulus is shown in Fig. 2. The plot includes the transient phase and shows how the periodic solution is approached. In particular, it
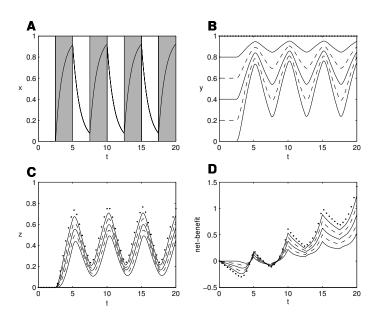

Figure 2: Simulation results in a periodic environment using the linear model. The plots show the response behaviour of x (A), y (B) and z (C) as well as the net-benefit of the system (D). Activating pulses are denoted by gray background. The expression levels k were set to $k = 0, 0.2, 0.4, 0.6, 0.8$ (solid and dashed, alternately from beneath) and $k = 1$ (dotted). The cost factors were set to $\eta_y = \eta_z = 0.1$.

is visible that higher basal expressions k leads to higher regulator concentrations y_k and to higher target expressions z_k. That still holds true even after the stimulating phases. This fading out can be interpreted as a memory of the biological system.

The investigation has shown that the system behavior typical for fluctuating environmental changes can be reproduced, but the linear model cannot provide a reason for any evolutionary advantage of a basal expression $k \in\ ]0, 1[$. Therefore, we need to investigate different model extensions.

4. Nonlinear model

A possible model extension of the initial model to reach a strong maximum net-benefit with $k \in\ ]0, 1[$ consists in using nonlinear activation functions $g_k(x)$ and $m(x, y)$ instead of the linear functions (10). We investigate system (1 - 3) with the aim of deducing an optimal net-benefit functional (7) with $k \in\ ]0, 1[$.

The saturated influence function

$$H(x) = \frac{(1 + K)x^h}{x^h + K} \tag{11}$$

with the saturation constant K and the Hill coefficient $h \geq 1$ fulfills $H(0) = 0$, $H(1) = 1$ and $\lim_{x \to \infty} H(x) = 1 + K$. Accordingly the half saturation of the function $H(x)$ is at $\bar{x} = \sqrt[h]{K}$.

We denote different saturated influence functions by indices

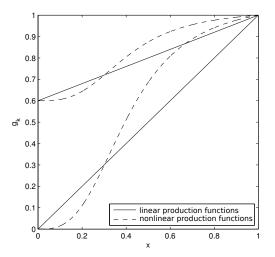$$H_{xy}(x) = \frac{(1 + K_{xy})x^h}{x^h + K_{xy}},$$

4

Figure 3: Production rates $g_k(x)$ of the second regulator y using the saturation function in Eq. (11) with $K = 0.07$ and $h = 3$ (dashed line) in comparison to the linear regulation in Eq. (10) with $k = 0$ and $k = 0.6$ (solid line).

where the substance scaled in x influences the production of the substance denoted by y. Of course, the saturation constant may depend on the involved substances. The Hill coefficient is generalized for all saturated influence functions.

The activation functions

$$g_k(x) = k + (1 - k)H_{xy}(x) \text{ and } m(x, y) = H_{xz}(x)H_{yz}(y) \quad (12)$$

obey the conditions in Eq. (5). In Figure 3 the linear and non-linear activation functions g_k are shown for two different basal expression levels ($k = 0$ and $k = 0.6$). The saturation branch with $H(x) > 1$ is not used in a realistic application of system (1 - 3) because $x \in [0, 1]$ as well as $y \in [0, 1]$ holds true for all t. Hence, the saturated influence function (11) is nonlinear for $x \in [0, 1]$ and a suitable parameter choice with $0 < K \ll 1$ approximates a saturation for values x near 1.

Now, we prove that a suitable parameter choice in the nonlinear saturated influence function can provide an optimum in the net-benefit functional (7) with $k \in]0, 1[$. So, we will see that it is possible to show an argument for the evolutionary advantage of a non-trivial basal expression in a nonlinear model.

Theorem 4.1.
There exist nonlinear influence functions (12) such that the model (1 - 3) has an optimal net-benefit $C(k)$ for $k \in]0, 1[$.

Proof. The notation y_k and z_k are again used to emphasize the dependence of the solution of system (1 - 3) on the basal expression k. We consider the derivative of the net-benefit function in (7)

$$C'(k) = \frac{1}{T} \int_0^T s_x \frac{\partial z_k}{\partial k} - \eta_y \frac{\partial g_k(x)}{\partial k} - \eta_z \frac{\partial m(x, y_k)}{\partial k} \, dt.$$

With the notations

$$\gamma_k(x) = \frac{\partial g_k(x)}{\partial k} \geq 0 \text{ and } m_y(x, y) = \frac{\partial m(x, y)}{\partial y}$$

and again with $y_k = L[g_k(x)]$ and $z_k = L[m(x, y_k)]$, the linearity of L provides

$$\frac{\partial y_k}{\partial k} = L[g_k(x)] \gamma_k, \quad \frac{\partial m(x, y_k)}{\partial k} = m_y(x, y_k) L[g_k(x)] \gamma_k$$

and

$$\frac{\partial z_k}{\partial k} = L[m(x, y_k)] m_y(x, y_k) L[g_k(x)] \gamma_k.$$

Now, we have the expression

$$C'(k) = \frac{1}{T} \int_0^T \gamma_k \Big((s_x L[m(x, y_k)] - \eta_z) m_y(x, y_k) L[g_k(x)] - \eta_y \Big) dt$$

or written as

$$C'(k) = \frac{1}{T} \int_0^T \gamma_k \Big((s_x z_k - \eta_z) y_k m_y(x, y_k) - \eta_y \Big) dt. \quad (13)$$

The influence function $m(x, y)$ can be chosen in a manner that $m_y(x, 1) = m_y(x, y_1) > 0$ is very small and that $m_y(x, y_0)$ is larger for all occurring values $y_0(t)$. If now η_z is small enough to assure a positive integral $s_x z_k - \eta_z$ for all k, a suitable η_y leads to

$$C'(0) > 0 \text{ and } C'(1) < 0$$

and therefore to the existence of a strong maximum of the net-benefit function $C(k)$ with $k \in]0, 1[$. □

The last argument for the choice of suitable parameters becomes more comprehensible when we consider the extreme scenario of a very fast oscillating external signal s_x. For $T_{i+1} - T_i \to 0$ for all i, the solution tends to constant functions $x \to \bar{s}$, $y_k \to g_k(\bar{s})$ and $z_k \to m(\bar{s}, g_k(\bar{s}))$. Then, the considered derivative of the net-benefit functional tends to

$$C'(k) \to \gamma_k(\bar{s}) \Big([\bar{s} \cdot m(\bar{s}, g_k(\bar{s})) - \eta_z] g_k(\bar{s}) m_y(\bar{s}, g_k(\bar{s})) - \eta_y \Big).$$
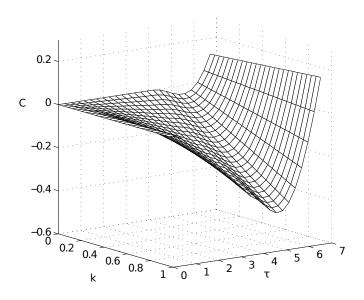
The choice $\eta_z < \bar{s} \cdot m(\bar{s}, g_0(\bar{s}))$ and a very small derivative $m_y(\bar{s}, 1) \ll m_y(\bar{s}, g_0(\bar{s}))$ near the full activation $y_1 = 1$ with

$$[\bar{s} m(\bar{s}, 1) - \eta_z] m_y(\bar{s}, 1) < [\bar{s} m(\bar{s}, g_0(\bar{s})) - \eta_z] g_0(\bar{s}) m_y(\bar{s}, g_0(\bar{s}))$$

allows us to choose η_y so that the assertion is fulfilled.

Just like the linear model, the nonlinear FFL model shows a negative instantaneous net-benefit c_k directly after the start of the first pulse at T_1. In Fig. 4 the net-benefit $C(k)$ is shown for changing basal expression levels k. The time after the first switch is defined as $\tau = t - T_1$.

Due to the availability of a strong maximum net-benefit with $k \in]0, 1[$, the model should be extended by maintaining the advantage of the basal expression and erasing the negativity of the instantaneous net-benefit. Therefore, we investigate the performance of the circuit and embed it in a basic model that mimics the energy flow within a cell in the next section.
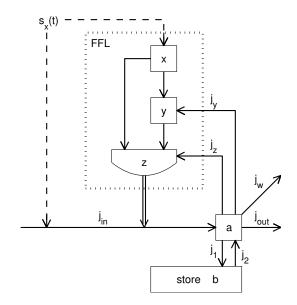
5

Figure 4: Net-benefit C in the nonlinear model, depending on the basal expression k and the time $\tau = t - T_1$ after the first switch. The cost factors are $\eta_y = \eta_z = 0.1$, the saturation constant is $K = 0.1$ and the Hill coefficient is $h = 3$. The net-benefit is negative for small τ. The instantaneous benefit rises above the instantaneous costs for larger τ and results in a positive net-benefit.

Figure 5: Expanded model with energy storing capacity in b and distributor a. The (x, y, z)-system is included as sub-model, marked by the dotted rectangle. Energy flows are marked by solid lines, signals by dashed lines, and the double arrow indicates the influence of z on the energy flow.

5. Model expansion by energy storing

So far we have concentrated on the net energy gain exclusively mediated by the regulatory circuit. This net energy gain has been identified with population growth. In particular, an instantaneous net energy loss leads to a decrease of the population size. However, in general, cells incorporate an energy storage system that allows for starvation without immediate death. For instance many bacteria, including *P. aeruginosa*, store energy in polyhydroxyalkanoates (PHAs) and

polyphosphates (PolyPs) to compensate for fluctuating environmental conditions (Chan et al., 2006; Achbergerová and Nahálka, 2011). In contrast, mainly ATP, but also other molecules as acetyl-CoA and NADH supply energy for immediate use. We embed our model in a larger structure addressing this energy storage system.

The storing system consists of two components: energy immediately available for usage $\tilde{a}$ ("distributor") and energy allocated to long-term store b ("store") (Langemann and Rehberg, 2010). Both components have some target size or setpoint $\tilde{a}_{\text{set}}$ resp. b_{set}. If $\tilde{a} < \tilde{a}_{\text{set}}$, there is not enough energy for maintenance and physiological processes available, and the cell starves. If $\tilde{a} > \tilde{a}_{\text{set}}$, the surplus of energy can be stored or used for reproduction. Similarly, $b < b_{\text{set}}$ may be detrimental, as time intervals without nutrient may lead to cell death; on the other hand, if $b > b_{\text{set}}$, energy is stored that could be used for reproduction. In our model, we do not address $\tilde{a}$ directly, but $a = \tilde{a} - \tilde{a}_{\text{set}}$, see Fig. 5. Note, that a may become negative. The considerations above indicate that the flow j_1 from a to b is proportional to $b_{set} - b$ (we allow a negative

flow here), and the flow j_2 from b to a is proportional to $a^- b^+$, where $a^- = (|a| - a)/2 = |\min(a, 0)|$ denotes the negative part of a, $a^+ = (|a| + a)/2 = \max(0, a)$ the positive part of a and $b^+ = (|b| + b)/2 = \max(0, b)$ denotes the positive part of b.

All energy flows enter or leave the central distributor a. The energy inflow $j_{\text{in}} = j_{\text{in}}(s_x, z)$ depends on the signal $s_x = s_x(t)$ and is modulated by the target z. We define $j_{\text{in}} = 1 - s_x(t) + s_x(t) z(t)$. A signal with activating stimulus $s_x(t) = 1$ provokes an energy flow mediated by z into the distributor a, cf. Eq. (4), and the last summand is non-vanishing. A signal without a stimulus $s_x(t) = 0$ affects an energy flow of size one, and the term $1 - s_x(t)$ is non-vanishing.

The circuit builds up proteins y and z at rates $g_k(x)$ resp. $m(x, y)$. As before, we assume that the energy required reads $j_y = \eta_y g_k(x)$ and $j_z = \eta_z m(x, y)$. To model the energy needed for general maintenance, we define an additional energy outflow j_{out} which is assumed to be constant in time.

The last, but important modeling step is to replace the net benefit C, see Eq. (6), by a compartment w that integrates all energy that is available for growth over time. Therefore, we define an energy flow j_w that is transformed into cell growth, i.e., from a to w. The cell starts to grow at an excess of immediately available energy, i. e. when $a > 0$ holds true. We define $j_w = \alpha a^+$.

6

Overall, the model equations read

$$
\begin{aligned}
\dot{x} &= s_x(t) - x, \\
\dot{y} &= g_k(x) - y, \\
\dot{z} &= m(x, y) - z, \\
\dot{a} &= j_{\text{in}} - j_1 + j_2 - j_y - j_z - j_{\text{out}} - j_w, \\
\dot{b} &= j_1 - j_2, \\
\dot{w} &= j_w.
\end{aligned}
\tag{14}
$$

Parameters that maximize w in the long run are considered to be optimal (Dekel et al., 2005). These parameters maximize the mean fitness

$$
W = \lim_{t \to \infty} \frac{w(t) - w(0)}{t}
\tag{15}
$$

The next proposition connects the basic model and its extended version.

The introduction of an energy-storing capacity interprets the increased energy need of the organism with basal expression in a modified way. This energy need is not longer seen as costs, but as the capability to use the stored energy faster.

It is clear, that the basal expression $k > 0$ is a cost in absence of signal s_x. Therefore, an organism with basal expression has lower remaining energy to grow in these phases. On the other hand, the energy in the store can be used faster if z is available.

Here, we demonstrate both effects by a periodic change of s_x with period $T = 5$, where the on phase $s_x(t) = 1$ and the off phase $s_x(t) = 0$ are of identical length. Furthermore, we use the linear influence functions in Eq. (10) and set

$$
b_{\text{set}} = 1, \quad j_1 = 1 - b, \quad j_2 = \beta a^- b^+
$$

with the parameters $\alpha = 4$, $j_{\text{out}} = 0.5$, $\eta_y = \eta_z = 0.1$. Fig. 6 shows the simulation results for different basal expressions k and a rather slow outflow $\beta = 1$ from the energy store.

The initial values are $x(0) = 0$, $y(0) = k$, $z(0) = 0$, $a(0) = 0$, $b(0) = b_{\text{set}} = 1$ and $w(0) = 0$.

As shown in Fig. 6B+C, a higher basal expression leads to higher production of transcription factor y and target z. In the first moment of adaptation, the energy of the store is used to produce both the y and z. In parallel, the energy in the distributor a and in the store b decreases while the fitness w stagnates especially for higher k values Fig. 6D-F.

This stagnation remains until the store is refilled. An organism with basal expression gets a considerate inflow j_{in} and is able to grow again earlier after the switch. In contrast, organisms with low basal expression need more time for adaptation, and they apparently grow faster for the first few cycles. This advantage is lost when the benefit of z raises the energy inflow j_{in} and organisms with high basal expression overtake all others. Even after four periods the organism with the highest basal expression has the most successful fitness in this example (Fig. 6F).

After the end of a signal, both the transcription factor and the target are still more abundant in organisms with basal expression. This can be considered as a sort of memory since
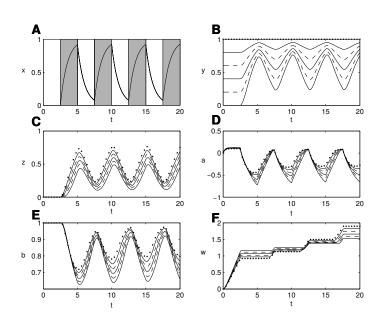

Figure 6: Simulation results in a periodic environment using the linear store model. In accordance with Fig. 2 the plots show the response behaviour of x (A), y (B) and z (C). The behaviour of distributor a and store b is shown in D and E, respectively. Increasing expression levels k lead to higher y, z, a, b while the fitness w is lower in the initial periods but overtakes after some cycles (F). Activating pulses are denoted by gray background. The expression levels k were set to $k = 0, 0.2, 0.4, 0.6, 0.8$ (solid and dashed, alternately from beneath) and $k = 1$ (dotted).

organisms are better prepared in an upcoming environmental change.

The role of the distributor in the model is similar to the one of the store, as it acts like a short time energy store. The role of the compartment a can be restricted to a pure distributor if the outflow j_2 from the store is increased under energy need by increasing β. In the limit situation $\beta \to \infty$, where the energy level in the distributor is constant, we have a distribution node without storing function. The system of ordinary differential equations transforms into a differential algebraic system with a constraint. On the other hand, the distributor can be seen as a distributing system in the organism which has a small storing function.

An extreme scenario with very fast switching for a model without a store is described in section 4, where $x \to \bar{s}$, $y_k \to g_k(\bar{s})$ and $z_k \to m(\bar{s}, g_k(\bar{s}))$ were found. For a very fast oscillating signal s_x the state variables a and b tend to constant functions. Consequently we get $j_1 = j_2$ and $j_w = j_{\text{in}} - j_y - j_z - j_{\text{out}}$. The growth $\dot{w} = j_w$ is calculated depending on $\bar{s}$ as

$$
\dot{w} = 1 - \bar{s} + \bar{s}(t)z - \eta_y g_k(\bar{s}) - \eta_z m(\bar{s}, g_k(\bar{s})) = \text{const.}
$$

The introduction of an energy storing system enables us to consider the initial production of the regulator y and the target gene z as an advantage for the organism. This can be seen in contrast to the previous sections 3 and 4 where the initial production was only interpreted as costs. Directly after the switch to the first pulse at $t = T_1 + \tau$, the growth $\dot{w} = j_w$ shows no negativity. The fitness w of the linear FFL model with an en-
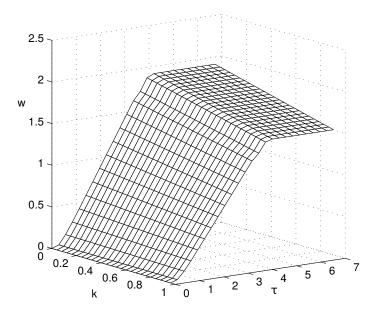
7

Figure 7: The fitness w of the linear model with an energy store system is shown for various basal expression levels k for times $\tau = t - T_1$ after the first switch T_1. The cost factors are $\eta_y = \eta_z = 0.1$ with $\alpha = 4, \beta = 10$ and $j_{\text{out}} = 0.5$. In this model, the fitness shows no negativity at all.

ergy store is shown in Figure 7 with cost factors $\eta_y = \eta_z = 0.1$, $\alpha = 4, \beta = 10$ and $j_{\text{out}} = 0.5$.

The use of linear influence functions (10) generates situations, where the long-time optimum of the fitness is at $k = 1$, see Fig. 6F. Since the long-time solution depends strongly monotonous on k, the linear model including a store is not suitable to explain the evolutionary advantage of a non-trivial basal expression.

In order to generate a model with the explanatory power including a store and the existence of a strong optimal basal expression discussed in Sec. 4, we combine both parts. Therefore, we now consider the system (14) with suitable nonlinear functions $g_k(x)$ and $m(x, y)$. We assume, that the nonlinear activation functions $g_k(x)$ and $m(x, y)$ are the same as in the nonlinear model without an energy store in Eq. 12.

The fitness w of the linear and the nonlinear store model occurs by different signal lengths and different basal expression levels. In the linear store model the maximal fitness w depending on the basal expression level k is found for minimal or for maximal basal expression. Using nonlinear activation functions different reaction times after environmental switching are expected. This is a sufficient condition for a strong maximal basal expression.

The model with the energy store requires the same optimal basal expression k as the model without an energy store, if the production functions are equal in both models. Finally the equality of the optimal basal expression k for the models with and without an energy store has to be proven.

Lemma 5.1. *We assume that $s_x(t)$ is periodic, and that the system (x, y, z, a, b) tends to a periodic solution. Then, a parameter*

set that optimizes model (14) also optimizes model (1)-(3) with cost functional (6) and vice versa.

Proof: We relate the mean fitness W in Eq. (15) and the net benefit C in Eq. (7) in the case of periodic solutions. We assume that the signal and the solution for model (14), apart from $w(t)$, are periodic with period T. In particular, this implies that the solution (1)-(3) is T-periodic, too. Integrating the differential equations over this period, we obtain

$$0 = \int_0^T \dot{b}(t)\, dt = \int_0^T j_1(t) - j_2(t)\, dt$$

from the equation for b and next, the equations for a and w provide

$$
\begin{aligned}
w(T) - w(0) &= \int_0^T j_w\, dt = \int_0^T j_{\text{in}} - j_y - j_z - j_{\text{out}}\, dt \\
&= \int_0^T 1 - s_x(t)\, dt - j_{\text{out}} T + T\, C \\
&= A + T\, C
\end{aligned}
$$

where we use

$$\int_0^T \dot{a}(\theta)\, d\theta = 0 \quad \text{and} \quad A = \int_0^T 1 - s_x(t)\, dt - j_{\text{out}} T.$$

Note that A does not depend on the model parameters. Hence, in each period, the gain of variable w equals C plus a constant. Thus, $W = A/T + C$. The parameters optimizing C also optimize W and vice versa. □

The models with and without an energy store lead to the same optimal basal expression. As shown in Lemma 5.1 the outlined energy store has no influence on the optimal basal expression. The optimal k leading to the optimized net-benefit C or to the optimized mean fitness W was determined for a wide range of periodic signals using the linear and nonlinear production functions. Simulations were carried out using the previously defined parameters at different signal lengths in a periods T. One period consists of a preceding time interval of length T_1 followed by the signal that lasts until the next switching point at time $T_2 = T$. Hereby, the ratio T_1/T gives the proportion of time in which the signal $s_x = 0$.

The linear model shows two stages. They are first, a full basal expression with $k = 1$ and second, a vanishing basal expression with $k = 0$. Depending on the environment the value of k switches between both levels, see Fig. 8A. Consequently, the linear production functions are not sufficient to explain the advantage of a non-trivial basal expression level $k \in\,]0, 1[$.

However, the model with nonlinear production functions show a continuous transition from zero to a full basal expression, see Fig. 8B. These models show a strong optimum and are able to explain the advantage of a FFL$_k$ type regulation. Remarkably, for longer periods (in simulation 8 at around $T > 12$)
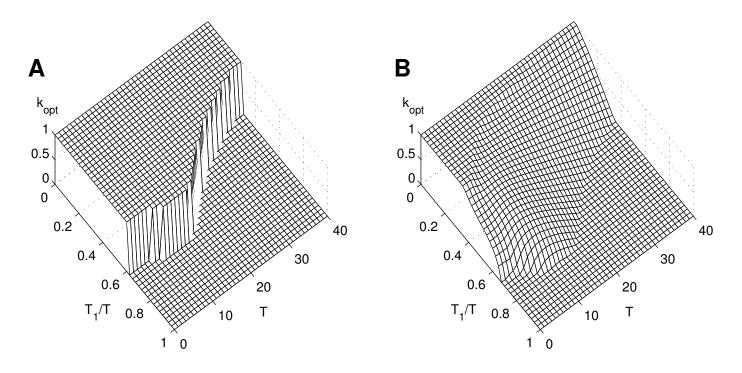
8

Figure 8: Optimal basal expression level k in various periodic environments determined by the period length T and the switching point T_1 which is the time point where the signal is turned on. In the linear model without a store or with a store (**A**) no optimal k exists while the nonlinear model with and without an energy store (**B**) shows a smooth region with $k \in]0, 1[$.

with sufficiently long parts without signal the same optimal basal expression value is found. This result is due to the fact that the state variables x, y and z decrease closely to their steady state level when the signal is lacking for longer time. However, long time-ranges without signal are a permanent cost factor in a FFL$_k$ regulatory circuit. Starting the signal from low-level state variables a higher basal expression is beneficial since the target is produced faster. In the case that the signal is not absent for a sufficiently long period of time, x, y and z are not completely degraded and the full benefit of a basal expression is not exploited. Here, parts of the induced compounds from the preceding stimulus are still provided, which can be considered as a kind of memory effect (Wolf et al., 2008).

6. Conclusion

From an abstract modeling viewpoint, we have investigated a model family of four mathematical models, which was generated by model extensions of Eqs. (1 -3) with linear production functions in section 3 in two directions. The first direction is the consideration of nonlinear activation functions in section 4, and the second direction is the inclusion of an energy storing compartment in section 5. The first extension enabled us to find an optimal basal expression k by using nonlinear production functions, and the second extension helped to reproduce a realistic behavior with a positive net-benefit, here redefined as growth, for organisms with FFL$_k$ type regulation. Each extension direction reproduces an observation from real biology. Finally, the model combining both extensions inherits the reproduction

of both observations. We can conclude, that the model extensions are robust with respect to each other and can be mapped to a certain observation from reality. The deterministic approach was chosen for the reason that the feed-forward system does not incorporate multi-stability. It is expected that stochastic effects merely lead to a fluctuation well centered around one mean value. In the long run, the fluctuations average out, and the efficiency of the system resembles in large parts that of the deterministic system considered here if the parameters are averaged in an appropriate way.

From the perspective of evolutionary biology FFL$_k$ type regulation produces costs in the absence of signal s_x. Thus, under this condition a FFL type regulation without a basal expression should be more beneficial with high selection pressure against FFL$_k$. Using the example of anaerobic adaptation in *P. aeruginosa*, the aerobic phases in a fluctuating environment cause permanent costs. However, the benefit arises during the shift to anaerobic conditions. In real life *P. aeruginosa* encounters mainly microaerobic and anaerobic conditions while living in the soil, on surfaces as biofilms or in the anaerobic mucus of cystic fibrosis patients (Schobert and Jahn, 2010). By applying the nonlinear store model it was shown that organisms with basal expression of the second regulator are better prepared to changing conditions. Hereby, the growth behaviour of the target genes z depends on the basal expression. The higher the basal expression is the earlier the production of the anaerobic system starts (Lemma 3.1). However, for a minimization of cost and a maximization of benefit a tradeoff between zero and maximum basal expression arises in fluctuating environments. The optimal basal expression k depends on the degree of environ-

9

mental fluctuation. In non-fluctuation environments no basal expression is required. On the other hand in extreme environments with nearly permanent s_x stimulus or under high fluctuation condition the optimum results in a full basal expression level. This state is equal to an independent simple-AND activation by both regulators (Dekel et al., 2005). A strong optimal basal expression exists only for pulses of sufficient length and frequency.

For this reason we argue that the anaerobic regulation system of *P. aeruginosa* is optimized to a balanced ratio of not too short anaerobic pulses. One of the main hazardous ways of living of *P. aeruginosa* is its persistence in biofilms e.g. in the airways of cystic fibrosis patients or in the urinary tract during its infections. It has been shown previously that the deeper layers of biofilms are anaerobic while close to the surface increasingly aerobic conditions are prevailing (Xu et al., 1998). In addition permanent reorganization of the biofilm structure especially in habitats with permanently flows, like urinary tract, leads to continuous adaptation demand (Lieleg et al., 2011). In this context FFL_k type regulation might be important for biofilm formation on surfaces and as a consequence for pathogenicity. The results also confirm that *P. aeruginosa* is a typical environmental bacterium adapted to many habitats rather than a highly specialized pathogenic organism adapted to the human host. This spectrum of ways of living is reflected by the versatile regulatory system employed by *P. aeruginosa*. Consequently, this mechanism of adaptation surely contributes to the evolutionary success of this group of bacteria.

There is increasing evidence that bacterial behavior strategies are often not just optimized for the actual conditions, but rather to cope with unpredictable fluctuating environments. Various strategies are known, including bet hedging and other forms of phenotypic heterogeneity, high mutation rates, or development of mobile genetic elements (Turrientes et al., 2013; Heuer et al., 2008; Salathé et al., 2009; Engelstädter and Moradigaravand, 2014; Gomes et al., 2013; Davidson and Surette, 2008). All of these include investment (costs) under certain conditions which pay off if the environment changes. Obviously, such strategies meet a central requirement for many species under natural conditions. From this perspective, it can be speculated that the FFL_k strategy may be more common, but sometimes overlooked as low basal gene expression is difficult to measure and could be misinterpreted in terms of unavoidable, but purposeless noise.

7. Acknowledgements

References

Achbergerová, L., Nahálka, J., 2011. Polyphosphate–an ancient energy source and active metabolic regulator. Microb. Cell. Fact. 10, 63.
URL http://dx.doi.org/10.1186/1475-2859-10-63

Alon, U., 2007. Network motifs: theory and experimental approaches. Nat. Rev. Genet. 8 (6), 450–461.
URL http://dx.doi.org/10.1038/nrg2102

Chan, P.-L., Yu, V., Wai, L., Yu, H.-F., 2006. Production of medium-chain-length polyhydroxyalkanoates by *Pseudomonas aeruginosa* with fatty acids and alternative carbon sources. Appl. Biochem. Biotechnol. 129-132, 933–941.

Davidson, C. J., Surette, M. G., 2008. Individuality in bacteria. Annu. Rev. Genet. 42, 253–268.
URL http://dx.doi.org/10.1146/annurev.genet.42.110807.091601

Dekel, E., Mangan, S., Alon, U., 2005. Environmental selection of the feed-forward loop circuit in gene-regulation networks. Phys. Biol. 2 (2), 81–88.
URL http://dx.doi.org/10.1088/1478-3975/2/2/001

Eichenberger, P., Fujita, M., Jensen, S. T., Conlon, E. M., Rudner, D. Z., Wang, S. T., Ferguson, C., Haga, K., Sato, T., Liu, J. S., Losick, R., 2004. The program of gene transcription for a single differentiating cell type during sporulation in *Bacillus subtilis*. PLoS Biol. 2 (10), e328.
URL http://dx.doi.org/10.1371/journal.pbio.0020328

Engelstädter, J., Moradigaravand, D., 2014. Adaptation through genetic time travel? fluctuating selection can drive the evolution of bacterial transformation. Proc. Biol. Sci. 281 (1775), 20132609.
URL http://dx.doi.org/10.1098/rspb.2013.2609

Gomes, A. L. C., Galagan, J. E., Segrè, D., 2013. Resource competition may lead to effective treatment of antibiotic resistant infections. PLoS One 8 (12), e80775.
URL http://dx.doi.org/10.1371/journal.pone.0080775

Görke, B., Stülke, J., 2008. Carbon catabolite repression in bacteria: many ways to make the most out of nutrients. Nat. Rev. Microbiology 6, 613–624.

Heuer, H., Abdo, Z., Smalla, K., 2008. Patchy distribution of flexible genetic elements in bacterial populations mediates robustness to environmental uncertainty. FEMS Microbiol. Ecol. 65 (3), 361–371.
URL http://dx.doi.org/10.1111/j.1574-6941.2008.00539.x

Kalir, S., Mangan, S., Alon, U., 2005. A coherent feed-forward loop with a SUM input function prolongs flagella expression in *Escherichia coli*. Mol. Syst. Biol. 1, 2005.0006.
URL http://dx.doi.org/10.1038/msb4100010

Lan, G., Tu, Y., 2013. The cost of sensitive response and accurate adaptation in networks with an incoherent type-1 feed-forward loop. J. R. Soc. Interface 10 (87), 20130489.
URL http://dx.doi.org/10.1098/rsif.2013.0489

Langemann, D., Rehberg, M., 2010. Unbuffered and buffered supply chains in human metabolism. J. Biol. Phys. 36 (3), 227–244.
URL http://dx.doi.org/10.1007/s10867-009-9178-4

Lieleg, O., Caldara, M., Baumgärtel, R., Ribbeck, K., 2011. Mechanical robustness of *Pseudomonas aeruginosa* biofilms. Soft. Matter. 7 (7), 3307–3314.
URL http://dx.doi.org/10.1039/c0sm01467b

Macia, J., Widder, S., Solé, R., 2009. Specialized or flexible feed-forward loop motifs: a question of topology. BMC Syst. Biol. 3, 84.
URL http://dx.doi.org/10.1186/1752-0509-3-84

Mangan, S., Alon, U., 2003. Structure and function of the feed-forward loop network motif. Proc. Natl. Acad. Sci. USA 100 (21), 11980–11985.
URL http://dx.doi.org/10.1073/pnas.2133841100

Mangan, S., Itzkovitz, S., Zaslaver, A., Alon, U., 2006. The incoherent feed-forward loop accelerates the response-time of the gal system of *Escherichia coli*. J. Mol. Biol. 356 (5), 1073–1081.
URL http://dx.doi.org/10.1016/j.jmb.2005.12.003

Mangan, S., Zaslaver, A., Alon, U., 2003. The coherent feedforward loop serves as a sign-sensitive delay element in transcription networks. J. Mol. Biol. 334 (2), 197–204.

Salathé, M., Van Cleve, J., Feldman, M. W., 2009. Evolution of stochastic switching rates in asymmetric fitness landscapes. Genetics 182 (4), 1159–1164.
URL http://dx.doi.org/10.1534/genetics.109.103333

Schobert, M., Jahn, D., 2010. Anaerobic physiology of *Pseudomonas aeruginosa* in the cystic fibrosis lung. Int. J. Med. Microbiol. 300 (8), 549–556.
URL http://dx.doi.org/10.1016/j.ijmm.2010.08.007

Schreiber, K., Krieger, R., Benkert, B., Eschbach, M., Arai, H., Schobert, M., Jahn, D., 2007. The anaerobic regulatory network required for *Pseudomonas aeruginosa* nitrate respiration. J. Bacteriol. 189 (11), 4310–4314.
URL `http://dx.doi.org/10.1128/JB.00240-07`

Shachrai, I., Zaslaver, A., Alon, U., Dekel, E., 2010. Cost of unneeded proteins in *E. coli* is reduced after several generations in exponential growth. Mol. Cell. 38 (5), 758–767.
URL `http://dx.doi.org/10.1016/j.molcel.2010.04.015`

Shen-Orr, S. S., Milo, R., Mangan, S., Alon, U., 2002. Network motifs in the transcriptional regulation network of *Escherichia coli*. Nat. Genet. 31 (1), 64–68.
URL `http://dx.doi.org/10.1038/ng881`

Silva-Rocha, R., de Lorenzo, V., 2011. A composite feed-forward loop I4-FFL involving IHF and Crc stabilizes expression of the XylR regulator of *Pseudomonas putida* mt-2 from growth phase perturbations. Mol. Biosyst. 7 (11), 2982–2990.
URL `http://dx.doi.org/10.1039/c1mb05264k`

Strohm, T. O., Griffin, B., Zumft, W. G., Schink, B., 2007. Growth yields in bacterial denitrification and nitrate ammonification. Appl. Environ. Microbiol. 73 (5), 1420–1424.
URL `http://dx.doi.org/10.1128/AEM.02508-06`

Turrientes, M.-C., Baquero, F., Levin, B. R., Martínez, J.-L., Ripoll, A., González-Alba, J.-M., Tobes, R., Manrique, M., Baquero, M.-R., Rodríguez-Domínguez, M.-J., Cantón, R., Galán, J.-C., 2013. Normal mutation rate variants arise in a mutator (mut s) *Escherichia coli* population. PLoS One 8 (9), e72963.
URL `http://dx.doi.org/10.1371/journal.pone.0072963`

Wall, M. E., Dunlop, M. J., Hlavacek, W. S., 2005. Multiple functions of a feed-forward-loop gene circuit. J. Mol. Biol. 349 (3), 501–514.
URL `http://dx.doi.org/10.1016/j.jmb.2005.04.022`

Widder, S., Solé, R., Macía, J., 2012. Evolvability of feed-forward loop architecture biases its abundance in transcription networks. BMC Syst. Biol. 6 (1), 7.
URL `http://dx.doi.org/10.1186/1752-0509-6-7`

Wolf, D. M., Fontaine-Bodin, L., Bischofs, I., Price, G., Keasling, J., Arkin, A. P., 2008. Memory in microbes: quantifying history-dependent behavior in a bacterium. PLoS One 3 (2), e1700.
URL `http://dx.doi.org/10.1371/journal.pone.0001700`

Xu, K. D., Stewart, P. S., Xia, F., Huang, C. T., McFeters, G. A., 1998. Spatial physiological heterogeneity in *Pseudomonas aeruginosa* biofilm is determined by oxygen availability. Appl. Environ. Microbiol. 64 (10), 4035–4039.