A Basic Mathematical Model for Iron Uptake in Bacteria

W.-J. Beyn
J. Rieger
*F*orschungsschwerpunkt Mathematisierung, Universität Bielefeld,
*P*ostfach 100131, Germany

1 Introduction

In this paper we present and discuss a first simple model for iron uptake and iron homeostasis in E.coli. The subject is interesting from the biological point of view, because some phenomena occuring in this context are not understood yet. Additionally, many bacteria have the ability to store iron in special iron storage clusters [1], which is very special and distinguishes iron from most other minor nutritients. On the other hand, the fact that iron is at the same time an indispensable nutritient and highly toxic for the cell indicates, that one might find a nontrivial, extraordinarily robust regulatory system. We were inspired by the research work [14, 2] at the Center of Biotechnology at the University of Bielefeld that investigates the symbiosis between S.meliloti and its host plant, where the iron supply of the symbiontic bacteroids plays a major role.

The mathematical modelling of intracellular processes is a central issue for several groups that work in this interdisciplinary field. For example, there are engineers and computer scientists [12] who cooperate with biologists and who predict by in-silico experiments which wet-lab experiment is likely to lead to the optimal gain of knowledge. Others develop theoretical concepts [11] in order to build tools which can assemble a model out of predefined subunits and simulate it. In most cases the systems are modelled by ordinary differential equations and they are solved numerically.

It is also possible to investigate a given system from a stoichiometric point of view. Often the systems are decomposed into so called elementary

flux modes [13] which can be used to optimize the yield of a desired product. Further attempts have been made to examine the stoichiometric robustness of a system against the elimination of components and draw conclusions for the evolution of intracellular networks [6].

This paper follows the first, more engineering-like approach. We state a simplistic ODE based model for iron uptake and iron homeostasis. In the discussion we will show its properties and limitations and propose improvements. We will see, that in the iron related context the stoichiometry of regulatory reactions is of an immense importance, and we will find that the regulation processes must be of an unusually high kinetic order.

In the future we intend to take a closer look at the subsystems and to model them on a higher detail level. Then we will try to use the relatively modern method of piecewise linear systems for genetic regulation ([4], for the mathematical background see [7]) that are not applicable for metabolic reactions, because they do not describe reactions of second order properly. Therefore we will investigate the properties of hybrid systems, which are piecewise linear as far as gene expression is concerned and classical ODEs in all other reactions.

2 Bacterial Iron Homeostasis from the Biological/Descriptive Point of View

Virtually all organisms are dependent on sufficient iron supply, because many reactions (e.g. in the TCA cycle or the respiratory chain) need it as a cofactor. The importance of iron results from its ability to adopt a redox potential in an exceptionally wide range from -300mV to +700mV depending on the surrounding conditions, which makes it a perfect electron acceptor and donor respectively.

On the other hand, free iron atoms in the cell are highly toxic. They catalyse reactions of reduced oxygen species, which are always present as a product of the energy metabolism, and convert them into radicals that can destroy the DNA. This means that organisms must have a very sophisticated regulation of the intracellular iron concentration.

Although iron is the fourth most plentiful element of the earth crust, the soluble form Fe^{2+} is scarce in the natural environment of bacteria. So they must have an active high affinity iron acquisition system which in gram

negative bacteria usually consists of receptors in the outer membrane and ABC-transporters in the cytoplasmatic membrane [3]. Both transport processes require energy provided by the cell. These transporters are directly inactivated by high intracellular iron concentrations, but the mechanism of this regulation is unknown. If the need for iron cannot be met by this method, many species excrete chelators called siderophores [10] which bind the unsoluble Fe^{3+} with a high affinity and try to reimport them.

Many bacteria produce iron storage proteins, and their structure and the way they work varies from species to species. The most prominent one is the bacterioferritin. The iron storage complex is a hollow sphere formed by 24 bacterioferritin subunits. It takes up the soluble Fe^{2+} in the cytoplasma and deposits it in its centre in the non-soluble form Fe^{3+}. The oxidisation may be carried out by a haem group bound to the bacterioferritin, but this is not entirely clear. Iron storage proteins are only expressed at low growth rates. This makes perfect sense in the natural environment of bacteria, e.g. in the soil. If some external event causes an increase of the concentration of nutritients and makes new growth possible, the species which already have stored iron have a big advantage compared to those which can just start to collect iron in the moment they sense their lack thereof.

In E.coli the global regulator for iron related processes is a protein called Fur (ferric uptake regulator). If it binds one Fe^{2+}-ion, it changes its conformation and dimerizes with another [Fur-Fe^{2+}]. This Fur-dimer can bind to a DNA sequence motif called the Fur-box and acts usually as a repressor by simply blocking the beginning of the operon. Induction of operons can be achieved by blocking a gene that again encodes a repressor RNA or a repressor protein. It has been shown in [1] that Fur-dimers do not simply bind to their box like other repressors do, but polymerize around the DNA in a corkscrew manner starting from the Fur box. The reason for this behaviour is still unclear.

The Fur modulon has been well studied. In the presence of iron, Fur represses the expression of the iron import machinery and induces the production of iron storage proteins and enzymes which need iron as a cofactor by derepression. This guarantees that energy and ressources are not wasted for the production of proteins which are unnecessary under those conditions. Fur has several other regulatory functions (see [9]) without direct impact on the subsystem of the cell we want to examine.

3 A basic model for iron homeostasis

The basic model features iron uptake, direct regulation of the transporters by the intracellular iron concentration, genetic regulation of the transport system mediated by the interior iron concentration, and the impact of iron as an important cofactor of enzymes on growth.

We want to model a classical experiment in stirred and spatially homogeneous medium. Therefore it is justified to denote all quantities as concentrations (mol/g dry weight, g/l for the biomass, and mol/l for the concentration of iron in the medium). The carbonhydrate source is assumed to be constant or rather so big that iron is the only growth limiting substrate.

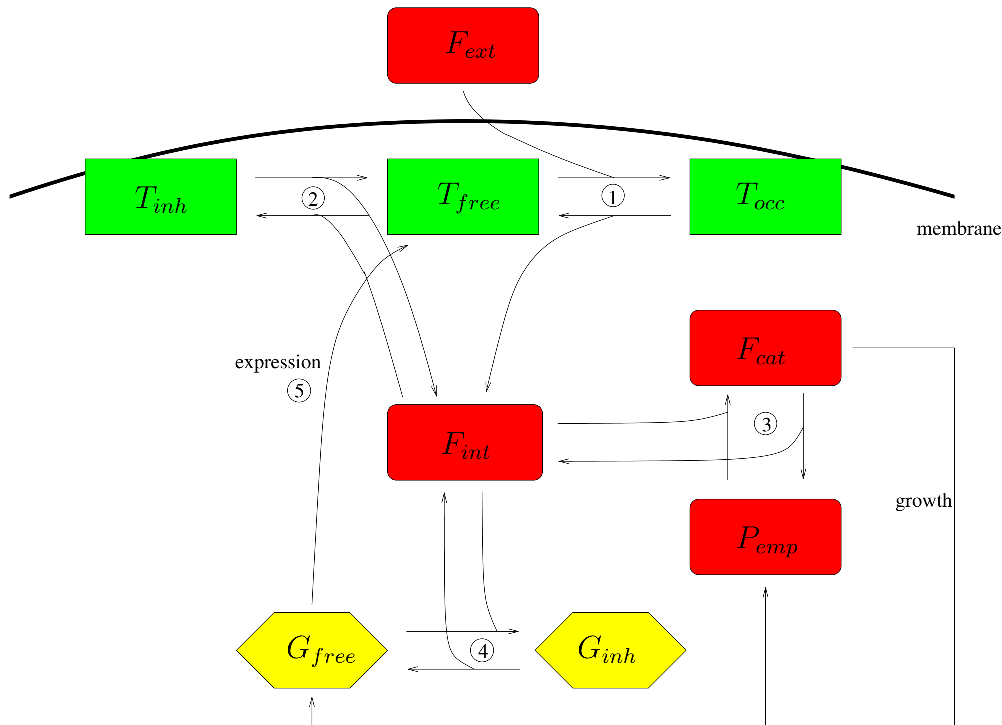


Abbildung 1: Basic model

3.1 Variables

The following variables occur in our model:

F_{ext}	–	extracellular iron,
F_{int}	–	free intracellular iron,
F_{cat}	–	intracellular iron which is bound in the catalytic center of enzymes,
P_{emp}	–	enzymes which need iron for their activity, ready to accept iron as cosubstrate,
T_{free}	–	free transporter, ready to accept extracellular iron,
T_{occ}	–	occupied transporter, already carrying an ion,
T_{inh}	–	inhibited transporter, blocked by a ligand dependent on intracellular iron,
G_{free}	–	an operon of genes encoding an ABC-transporter, ready to be transcribed and translated,
G_{inh}	–	an operon of genes encoding an ABC-transporter, blocked by an iron-dependent repressor,
M	–	biomass in the medium.

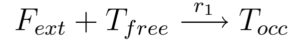
3.2 Constants

The constants below define a framework for the simulation. All other constants are rates of the reactions described later.

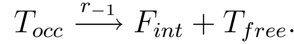
μ_{max}	–	maximal growth rate under fully iron saturated conditions,
μ_s	–	concentration of the limiting substrate (in this case the surplus of catalytically active iron which is not required for the bare survival of the cell) at which the bacteria grow with half their maximal rate,
β_{fe}	–	concentration of catalytically active iron [mol l^{-1}] required by unit of biomass [gl^{-1}] for bare survival,
γ	–	concentration of iron dependent enzymes [mol l^{-1}] per unit of biomass [gl^{-1}],
ω	–	number of cells [mol l^{-1}] per unit of biomass [gl^{-1}],
F_{all}	–	total concentration of iron in the system.

3.3 Reactions

1. iron uptake and release into cytoplasma by a transporter

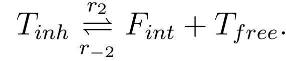


$$F_{ext} + T_{free} \xrightarrow{r_1} T_{occ}$$



$$T_{occ} \xrightarrow{r_{-1}} F_{int} + T_{free}.$$

2. direct downregulation of transporter activity by intracellular iron concentration



$$T_{inh} \underset{r_{-2}}{\overset{r_2}{\rightleftharpoons}} F_{int} + T_{free}.$$

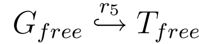
3. insertion of free iron into enzymes

$$P_{emp} + F_{int} \underset{r_{-3}}{\overset{r_3}{\rightleftharpoons}} F_{cat}.$$

4. regulation of transporter gene transcription

$$G_{free} + F_{int} \underset{r_{-4}}{\overset{r_4}{\rightleftharpoons}} G_{inh}.$$

Further we model the expression of the transporter genes

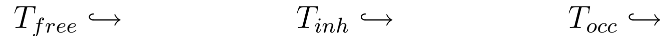


$$G_{free} \overset{r_5}{\hookrightarrow} T_{free}$$

cell growth dependent on catalytically active iron

$$F_{cat} \rightsquigarrow M$$

and the decay of the various types of transporters



$$T_{free} \hookrightarrow \qquad T_{inh} \hookrightarrow \qquad T_{occ} \hookrightarrow$$

If a transporter which carries an ion (T_{inh}, T_{occ}) decays, we assume that the iron is released into the cytoplasma. The amount of iron accepting enzymes (in total $P_{emp} + F_{cat}$) is considered as a constant part of the biomass M and therefore to be directly controlled by cell growth. The cell growth is modelled by an equation for one limiting substrate similar to Monod and Michaelis-Menten kinetics.

All concentrations apart from F_{ext} have to be scaled against the growth of biomass. Therefore we add an extra term of the form $-S\frac{\dot{M}}{M}$ to each equation for the concentration S [mol/g].

3.4 The system

Now we can state the model as a ten-dimensional system of ordinary differential equations with eleven parameters:

$$
\begin{aligned}
\dot{F}_{ext} &= -r_1 F_{ext} T_{free} M \\
\dot{F}_{int} &= r_{-1} T_{occ} - r_{-2} F_{int} T_{free} + r_2 T_{inh} - r_4 F_{int} G_{free} + r_{-4} G_{inh} \\
&\quad - r_3 P_{emp} F_{int} + r_{-3} F_{cat} + d_T (T_{occ} + T_{inh}) - F_{int} \frac{\dot{M}}{M} \\
\dot{F}_{cat} &= r_3 P_{emp} F_{int} - r_{-3} F_{cat} - F_{cat} \frac{\dot{M}}{M} \\
\dot{P}_{emp} &= -r_3 P_{emp} F_{int} + r_{-3} F_{cat} + F_{cat} \frac{\dot{M}}{M} \\
\dot{T}_{free} &= r_5 G_{free} + r_{-1} T_{occ} - r_1 F_{ext} T_{free} + r_2 T_{inh} - r_{-2} F_{int} T_{free} \\
&\quad - d_T T_{free} - T_{free} \frac{\dot{M}}{M} \\
\dot{T}_{occ} &= r_1 F_{ext} T_{free} - r_{-1} T_{occ} - d_T T_{occ} - T_{occ} \frac{\dot{M}}{M} \\
\dot{T}_{inh} &= r_{-2} F_{int} T_{free} - r_2 T_{inh} - d_T T_{inh} - T_{inh} \frac{\dot{M}}{M} \\
\dot{G}_{free} &= -r_4 F_{int} G_{free} + r_{-4} G_{inh} + G_{inh} \frac{\dot{M}}{M} \\
\dot{G}_{inh} &= r_4 F_{int} G_{free} - r_{-4} G_{inh} - G_{inh} \frac{\dot{M}}{M} \\
\dot{M} &= \frac{\mu_{max} \left(F_{cat} - \beta_{fe}\right)}{\mu_s + \left(F_{cat} - \beta_{fe}\right)} M
\end{aligned}
$$

3.5 Justification of the model

Due to the lack of solid measurable data, of detailed insight into the iron metabolism of the organism, and in particular of reliable kinetic parameters it does not make sense to create a big model in the first approach. We want to learn more about the intrinsic conditions of cellular systems by studying this first simple model and its limitations. The whole iron uptake process has been simplified drastically. In this first approach, we regard it as a black box binding extracellular iron and releasing it into the cytoplasma. As the knowledge about the effect of iron on growth is limited, we use a standard

differential equation, which describes growth on a limiting substrate. It is well known that the expression of many enzymes which use iron as a cofactor is stimulated by iron and negatively autoregulated. The latter implies that there exists an upper bound for the relative concentration of enzyme-bound iron. Due to the fact that there is little knowledge about the kinetics of the expression of these enzymes we regard it as a reaction in steady state, which means that the concentration of these enzymes always achieves its upper bound. The expression of the transporter genes is assumed to be regulated only by a global iron binding repressor. We omit the explicit modelling of the regulators and use a kind of book-keeping principle instead. View for example the regulation of the transporter expression. The free operon G_{free} and the free internal iron F_{int} form a complex which is the repressed operon G_{inh} containing the iron that is in reality bound by the global repressor. We have completely left out the modelling of iron storages, because these molecules are usually only expressed when the carbohydrate source is consumed and there is still iron in the medium, which is never the case in this artificial experimental setup.

3.6 Reduction of the model

The model describes an experiment in a closed system (e.g. a shake flask). Hence the total concentration of iron cannot vary. Therefore we have the conservation law

$$F_{all} = F_{ext} + M(F_{int} + F_{cat} + T_{occ} + T_{inh} + G_{inh})$$

where F_{all} is a constant. As we assume that the cells are more or less of the same size, we can relate the biomass to the total number of operons encoding the ABC-Transporter by

$$\omega = G_{free} + G_{inh}.$$

As described above, the iron binding enzymes are assumed to be a constant part of the biomass. Therefore

$$\gamma = P_{emp} + F_{cat}.$$

Now we can eliminate the three variables F_{ext}, P_{emp}, and G_{inh} by substituting

$$\begin{aligned} G_{inh} &= \omega - G_{free}, \\ F_{ext} &= F_{all} - M(F_{int} + F_{cat} + T_{occ} + T_{inh} + G_{inh}) \\ &= F_{all} - M(F_{int} + F_{cat} + T_{occ} + T_{inh} + \omega - G_{free}), \\ P_{emp} &= \gamma - F_{cat}. \end{aligned}$$

4 Simulations

The numerical computations were tougher than we had expected. The system still has to be scaled. We integrated the system with the radau5 solver using extremely crude error tolerances, but even this procedure worked just for a small selection of physically relevant initial values. We think these problems are due to the fact that we are dealing with variables and parameters of very different order of magnitude. A. Kremling from the Max-Planck-Institut at Magdeburg reported that this difficulty is common. The inconveniency might be moderated by a nondimensionalisation, which we postpone, because recent calculations show that this model, derived from standard kinetic laws, is not sufficiently accurate (see section 4.3.).

4.1 Constants, Parameters, and Initial Values

In the model there are some constants which can be determined easily. We defined ω as the number of cells [mol l^{-1}] per unit of biomass [gl^{-1}] in dry weight. According to E. Flaschel, one cell of E.coli has a dry weight of $2,8 \cdot 10^{-13}$g. Hence a concentration of 1gl^{-1} implies a density of cells of $3.6 \cdot 10^{12} l^{-1}$, and this equals $6 \cdot 10^{-12}$mol l^{-1}. The constant γ symbolizes the concentration of iron dependent enzymes per unit of biomass [gl^{-1}]. According to S. Andrews [1] the concentration F_{cat} never exceeds $10^6 \omega M$ and never drops lower than $10^5 \omega M$. We think that the cell does not waste energy for the production of enzymes which are never used, and assume that under iron replete conditions 90 percent of the enzymes should contain all the iron they need. This leads to $\gamma = \frac{100}{90} 10^6 \omega = 6.67 \cdot 10^{-6}$. The lower bound for F_{cat} is the minimum concentration a cell can cope with, and so by definition $\beta_{fe} = 10^5 \omega = 6 \cdot 10^{-7}$. The minimal replication time of 20 minutes for E. coli can be found in every standard book on microbiology. This means that the

maximal growth rate can be determined as follows:

$$e^{\mu_{max}\cdot 20\text{min}} = 2 \Longrightarrow \mu_{max} = \frac{\ln 2}{20\text{min}} = 3.466\cdot 10^{-2}\text{min}^{-1}.$$

We developped a nice method which allows to find suitable kinetic parameters and initial values for the numerical integration at once: Regard the system in the steady state, which is defined by assuming that all (fast) single reactions are in an equilibrium. Then try to calculate by algebraic means, if the biological data can be realized in this or a slightly perturbed state. The main idea is that if one has an upper and a lower bound for at least one intracellular concentration, these bounds define a physiologically relevant interval $[S^{min}, S^{max}]$. This interval can be mapped monotonously to physiologically relevant intervals $[V^{min}, V^{max}]$ for other concentrations V simply by solving the steady state equations, because reactions in biological systems are always monotonous in the concentrations of their substrates. If one has more than one substrate for which upper and lower bounds are known one must take the intersections of the resulting relevant intervals $\cap_k [V_k^{min}, V_k^{max}]$. If the model represents the biological reality, it must be possible to find suitable parameters by algebraic calculations such that all measured data is contained in or is very close to these intervals. Hence comparison of the intervals and the data gives rise to constraints on the parameters, which are easy to compute.

The problem is how to define *very close* and *slightly perturbed*. As there are usually many parameters which cannot be estimated, i.e. more degrees of freedom than one can handle, we propose not to allow any perturbation and to insist that all the data must be contained in the intervals. The following calculations illustrate this concept.

In our special case, we look at the steady state the bacteria are in before they are put into the medium. The steady state condition implies that the cells have used up all their iron for growth and F_{cat} has achieved its minimum level. We denote all concentrations in this state of minimal iron supply by S^{min}. In the end, these concentrations will be suitable initial values. As explained above, it is useful to look at the system under fully iron saturated conditions, too. In this state we denote the concentrations by S^{max}.

We start our experiment with 10^6 cells per litre which is $1.67\cdot 10^{-18}$mol l^{-1} or $2.9\cdot 10^{-7} gl^{-1}$. The steady state condition implies

$$\dot{M} = 0$$

and therefore

$$F_{cat}^{min} = 10^5 \omega M = 1.68\cdot 10^{-13}.$$

Initial Conditions (X^{min})	
F_{ext}	10^{-5}
F_{int}	$5.68 \cdot 10^{-18}$
F_{cat}	$1.68 \cdot 10^{-13}$
P_{emp}	$1.7 \cdot 10^{-12}$
T_{free}	$6.54 \cdot 10^{-15}$
T_{occ}	0
T_{inh}	$7.34 \cdot 10^{-14}$
G_{free}	$7.99 \cdot 10^{-19}$
G_{inh}	$8.81 \cdot 10^{-19}$
M	$2.8 \cdot 10^{-7}$

In the iron replete case (s.a.) we have

$$F_{cat}^{max} = 10^6 \omega M = 1.68 \cdot 10^{-12}.$$

According to the conservation law

$$P_{emp} = \gamma M - F_{cat}$$

we obtain

$$\begin{aligned} P_{emp}^{max} &= \gamma M - F_{cat}^{max} = 1.87 \cdot 10^{-12} - 1,68 \cdot 10^{-12} = 1,2 \cdot 10^{-13}, \\ P_{emp}^{min} &= \gamma M - F_{cat}^{min} = 1.87 \cdot 10^{-12} - 1,68 \cdot 10^{-13} = 1,7 \cdot 10^{-12}. \end{aligned}$$

We know from experimental data [1] that the concentration of free intracellular iron is about 300 atoms per cell. It was impossible to measure an upper or lower threshold for this variable. Hence we have

$$F_{int} \approx 300 \omega M = 5.04 * 10^{-16}.$$

Now we try to find good kinetic parameters r_3, r_{-3}. We focus only on the reaction $F_{cat} \leftrightharpoons F_{int} + P_{emp}$:

$$\begin{aligned} 0 = \dot{F_{cat}} &= r_3 P_{emp} F_{int} - r_{-3} F_{cat} \\ &= r_3 (\gamma M - F_{cat}) F_{int} - r_{-3} F_{cat} \end{aligned}$$

and therefore

$$\begin{aligned} F_{int}^{max} &= \frac{r_{-3}F_{cat}^{max}}{r_3(\gamma M - F_{cat}^{max})} = 8.84\frac{r_{-3}}{r_3}, \\ F_{int}^{min} &= \frac{r_{-3}F_{cat}^{min}}{r_3(\gamma M - F_{cat}^{min})} = 9.87 \cdot 10^{-2}\frac{r_{-3}}{r_3}. \end{aligned}$$

Determine $R := \frac{r_{-3}}{r_3}$ such that the sum of the errors

$$(F_{int}^{max} - 5.04 \cdot 10^{-16})^2 + (F_{int}^{min} - 5.04 \cdot 10^{-16})^2$$

achieves its minimum. As this defines a real quadratic function, the following zero of the derivative is the global minimum of the upper formula:

$$\begin{aligned} 0 &= 2 \cdot 8.84(8.84R - 5.04 \cdot 10^{-16}) \\ &\quad +2 \cdot 9.87 * 10^{-2}(9.87 \cdot 10^{-2}R - 5.04 \cdot 10^{-16}) \\ \Longleftrightarrow &\quad (9.74 \cdot 10^{-3} + 78.146)R = 4.5 \cdot 10^{-15} \\ \Longleftrightarrow &\quad R = 5.76 \cdot 10^{-17}. \end{aligned}$$

Thus

$$\begin{aligned} F_{int}^{max} &= 5.09 \cdot 10^{-16} \approx 303 \text{ ions per cell} \\ F_{int}^{min} &= 5.68 \cdot 10^{-18} \approx 3 \text{ ions per cell} \end{aligned}$$

This approach strongly emphasizes the upper bound, which is appropriate, because too high concentrations of intracellular free iron cause cell death.

We do not have any information about the affinity of Fur to the Fur box or estimates for the percentage of repressed operons under defined conditions. Hence we cannot avoid to use intuitively plausible estimates of the percentage of active ABC-transporter encoding operons:

$$G_{free}^{max} = \frac{1}{100}\omega M,$$

$$G_{free}^{min} = \frac{90}{100}\omega M.$$

The really important bound is the downregulation by high internal iron concentrations. Therefore we fit the kinetic rates to this value:

$$\begin{aligned} 0 = \dot{G_{inh}} &= r_4 F_{int} G_{free} - r_{-4} G_{inh} \\ &= r_4 F_{int} G_{free} - r_{-4}(\omega M - G_{free},) \end{aligned}$$

Parameters and Constants (time in minutes)	
ω	$6 \cdot 10^{-12}$
γ	$6.67 \cdot 10^{-6}$
β_{fe}	$6 \cdot 10^{-7}$
μ_{max}	$3.46 \cdot 10^{-2}$
μ_s	10^{-5}
r_1	$6 \cdot 10^4$
r_{-1}	$5 \cdot 10^4$
r_2	$6 \cdot 10^3$
r_{-2}	$1.18 \cdot 10^{22}$
r_3	$1.04 \cdot 10^{19}$
r_{-3}	600
r_4	$1.16 \cdot 10^{19}$
r_{-4}	60
r_5	$2 \cdot 10^5$
d_T	2

$$\iff \qquad G_{free} = \frac{r_{-4}\omega M}{r_4 F_{int} + r_{-4}},$$

and

$$\frac{1}{100}\omega M = G_{free}^{max} = \frac{r_{-4}\omega M}{r_4 F_{int}^{max} + r_{-4}}$$

$$\implies r_4 = \frac{99}{F_{int}^{max}} r_{-4} = \frac{99}{5.09 \cdot 10^{-16}} r_{-4} = 1.94 \cdot 10^{17} \cdot r_{-4}.$$

This means that under low iron conditions we have

$$G_{free}^{min} = \frac{r_{-4}\omega M}{r_4 F_{int}^{min} + r_{-4}} = 7.99 \cdot 10^{-19} \approx \frac{48}{100}\omega M$$

which is certainly too low, because nearly all operons should be active now. The conservation law for DNA gives

$$G_{inh}^{min} = \omega M - G_{free}^{min} = 1.68 \cdot 10^{-18} - 7.99 \cdot 10^{-19} = 8.81 \cdot 10^{-19}$$

According to [1] an average E.coli cell carries 1000 transporters when saturated with iron, and up to 100000 transporters under low iron conditions.

We denote the total concentration of transporters by T_{all}. Then we have

$$0 = \dot{T_{all}} = r_5 G_{free} - d_T T_{all}.$$

To be consistent, we look at the iron replete case for the fitting of the regulatory reactions:

$$1000 \cdot \omega M = T_{all}^{max} = \frac{r_5}{d_T} G_{free}^{max} = \frac{r_5}{d_T} \frac{1}{100} \omega M,$$

thus

$$r_5 = 10^5 d_T$$

This means for iron starvation conditions

$$T_{all}^{min} = \frac{r_5}{d_T} G_{free}^{min} = 10^5 \cdot 7.99 \cdot 10^{-19} = 7.00 \cdot 10^{-14} \approx 47560 \text{ transporters per cell}$$

This value is again definitely too low.

Finally we have to treat the direct regulation of transporters by the interior iron concentration. We do not have experimental data here, but it is clear that this regulation must be extremely strict. We assume that under iron replete conditions, only one transporter per cell can be open. This means that, for $T_{open} = T_{all} - T_{inh}$

$$\begin{aligned} 0 = \dot{T_{open}} &= r_2 T_{inh} - r_{-2} F_{int} T_{open} \\ &= r_2 (T_{all} - T_{open}) - r_{-2} F_{int} T_{open} \end{aligned}$$

$$\Longrightarrow \qquad T_{open} = \frac{r_2 T_{all}}{r_2 + r_{-2} F_{int}}$$

and hence

$$\begin{aligned} 0 &= r_2 (T_{all}^{max} - T_{open}^{max}) - r_{-2} F_{int}^{max} T_{open}^{max} \\ &= r_2 (1000 \cdot \omega M - \omega M) - 5.09 \cdot 10^{-16} r_{-2} \omega M \\ \Longrightarrow r_2 &= 5.065 \cdot 10^{-19} r_{-2}. \end{aligned}$$

When iron is scarce we find

$$T_{open}^{min} = \frac{r_2 T_{all}^{min}}{r_2 + r_{-2} F_{int}^{min}} = 6,54 * 10^{-15} \approx 3893 \text{ active transporters per cell.}$$

This is a very bad number. Nearly all transporters should be open now. As $T_{occ} = 0$ in the equilibrium before the cultivation, we must have

$$T_{free}^{min} = T_{open}$$

and

$$T_{inh}^{min} = T_{all}^{min} - T_{open}^{min} = 7.34 \cdot 10^{-14}.$$

Due to the fact that we determined the parameters by conditions on the steady state we just found ratios of pairs of parameters. We obtain the real parameters by guessing a value of the right order of magnitude resp. timescale for one of each pair. Both r_1 and r_{-1} are clearly on the millisecond timescale and can neither be calculated nor found in literature. The mechanism of the direct regulation of the transporters is still unknown. It makes sense to assume that it is not much slower than the iron import (in the case of siderophore transprot it might even be quicker). The insertion of iron into enzymes should be considerably slower, because iron is not inserted as a single ion, but as a bigger component such as a haem group. We estimate the movement of the repressor protein to be on the second and the decay of the transporter proteins on the minute time scale. Choose μ_s such that the duration of the cultivation is appropriate.

4.2 Numerical Simulation

As mentioned before we had problems with the integration of the data. For this reason we had to introduce two modifications: The expression for the growth rate (see for example the 10th component of the full system) has a singularity at $-\mu_s$. The biomass M can never reach or cross this unbiological value, but the numerical solver might have to use this point or points beyond for the calculation of a suitable stepsize. As μ_s is very small, this is likely to happen. So we cut off the expression for the growth rate on the left of $-\frac{1}{2}\mu_s$ and extended it by a straight line without changing the physiologically relevant part on the right of $-\frac{1}{2}\mu_s$.

The equation was still only badly solvable. So we searched for a set of initial values and parameters very close to the original ones which lead to better results. This slight perturbation does not really matter, because we have some uncertainties in these values anyway. We have used the program 'content' for the creation of graphs of the functions. The results are the following:

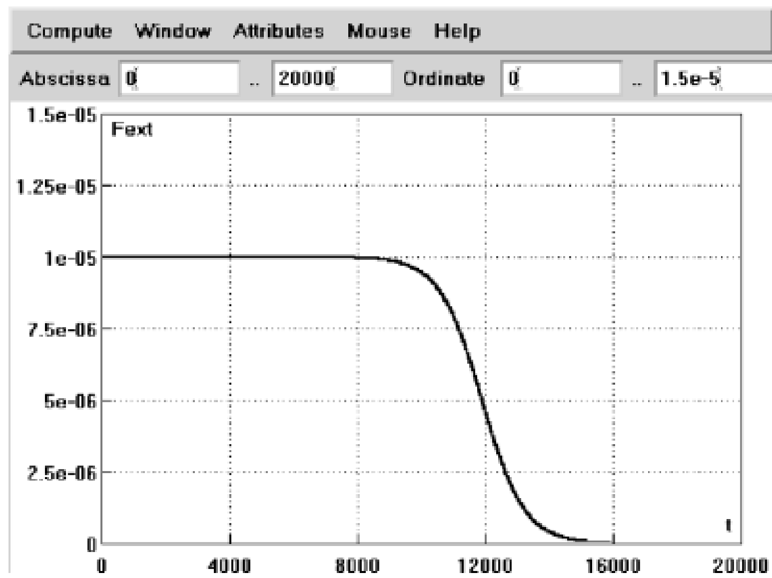


Abbildung 2: consumption of external iron

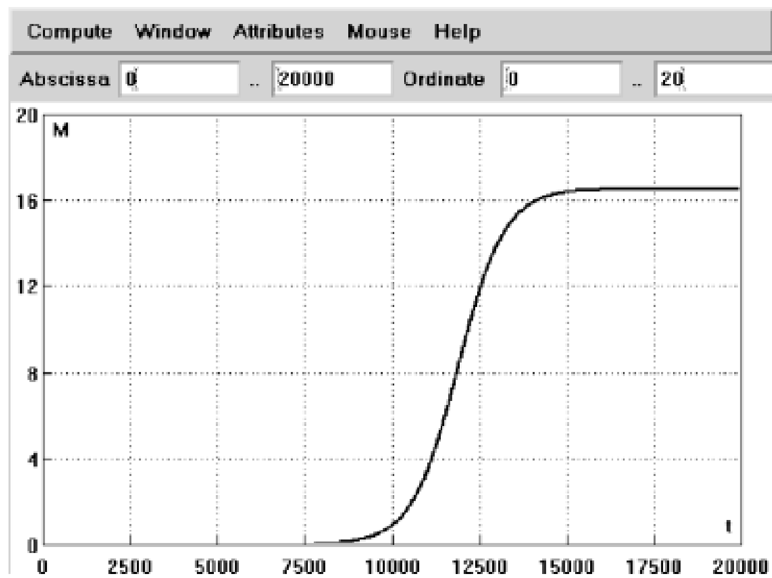


Abbildung 3: biomass

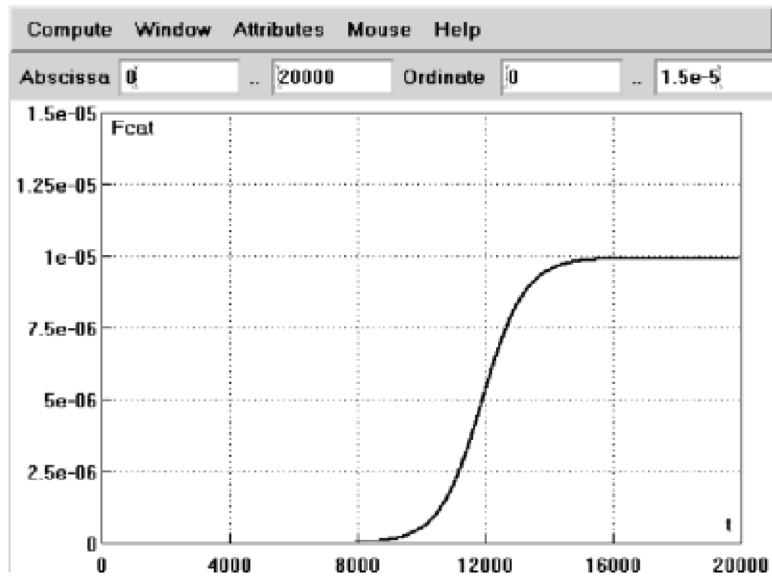


Abbildung 4: increase of catalytically active iron

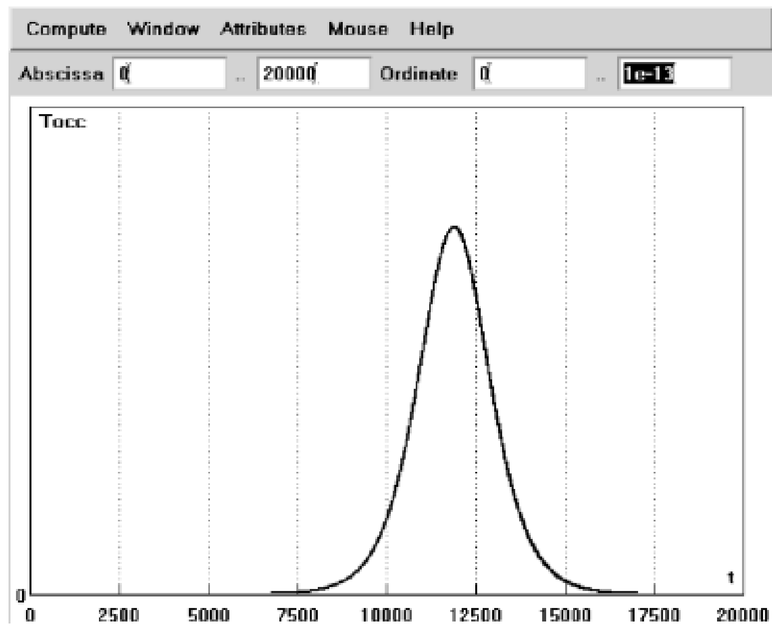


Abbildung 5: busy transporters

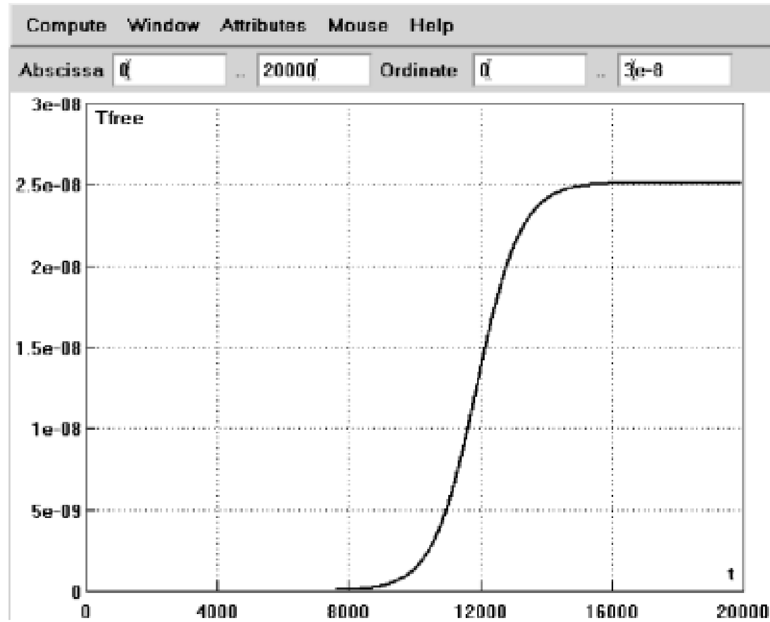


Abbildung 6: free transporters

The curves look exactly as they should. We do not have enough experimental data to calculate how good they are, but they display the right qualitative behaviour. Although it was very difficult to find initial values for which numerical integration was possible, we made some tests to check the robustness of the system. After some incubation time we disturbed it by suddenly changing an internal concentration. The system behaved exactly as one imagines, gliding swiftly into its internal equilibrium. This is not really surprising, because we have three feed back loops which contribute to keeping the concentrations well-balanced.

4.3 Discussion of the Simulations

As we have seen in the calculations from which we obtained our set of parameters and initial conditions, the structure of the model cannot describe the biological situation appropriately. We described the regulatory processes as single reactions, where one ion and one active unit form an inactive complex which can dissociate again. The trouble originates from the fact that there is an upper and a lower bound for the concentration of F_{cat} in the living cell. By

the steady state condition this physiologically relevant interval $[F_{cat}^{min}, F_{cat}^{max}]$ is mapped more or less linearly to relevant intervals $[X^{min}, X^{max}]$ for the other concentrations, which often don't agree with intelligent guesses and the little experimental data we have. The calculations show that we cannot fit the intervals $[X^{min}, X^{max}]$ to the real data with such a simple model. Very bad values occur in the number of active transporters where we see a cumulative effect of suboptimal genetic regulation and a direct downregulation of the transporters, which is by far too weak.

In order to solve these problems, we propose to pursue a more detailed modelling of the biological phenomena. There must be highly nonlinear effects in the regulatory reactions we have not considered so far. In the introduction we have mentioned the polymerisation of the repressor around the DNA helix, and we think that this polymerisation process is the source of nonlinearity which is responsible for a high order of the downregulation.

The mechanism for the direct inactivation of the ABC-transporters is still unknown, but our model shows that it should be of an even higher order, because it must be extremely strict under iron replete conditions and very relaxed when there is only few iron in the cell. For that reason it is necessary that at least two and perhaps more ions at once are involved in blocking one transporter. One can only speculate if they bind as ions or if they are associated with one or more components like proteins or iron-sulfur-clusters. If iron would bind to the transporters as single ions the reaction on instreaming iron would be very fast and direct, but it is hard to imagine a protein that is gradually inactivated. An iron-protein complex diffuses very slowly and is therefore no suitable signal for this delicate regulation. Thus it would make sense to expect iron binding components like the ubiquitous iron-sulfur clusters.

Maybe the high kinetic order of the transporter inactivation is in reality a spatial phenomenon. As the imported iron is localized near the cell membrane for at least a short period of time, it is possible that the interaction of free iron with other molecules is much stronger there than in the center of the cell, because the ions are not homogenuously distributed. In order to check this, one would have to integrate a reaction-diffusion system inside the cell, which is very different from the standard reaction-diffusion systems in pure liquid, because the cytoplasma is so packed with biomolecules.

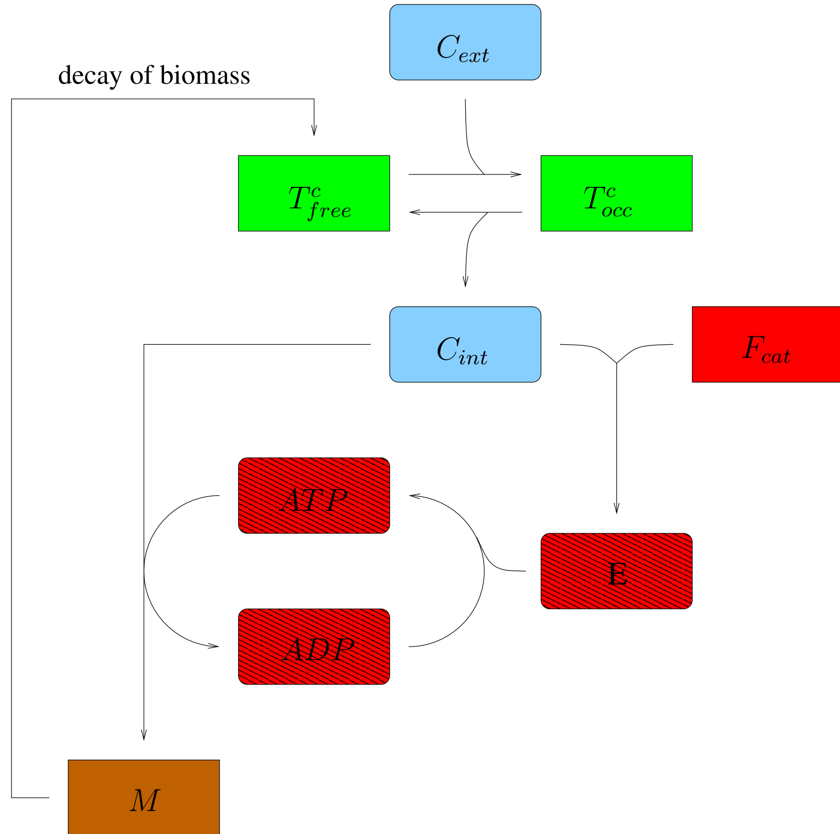


Abbildung 7: growth dependent on c-source and iron

5 Extensions of the Model

5.1 Structural Extensions

Once we have optimized the small model we discussed here we would like to describe the iron homeostasis in a much more precise way. There are at least two important points that should be improved: The growth of the biomass should be modelled as the influx of a carbohydrate source minus the total energy consumption. Before we can do that we need a realistic estimate for the stochiometry of these fluxes. This is hard to determine, because we are summing up many reactions as one. We will have to ask for the biologists' advice again and then make a good guess for suitable coefficients. We will

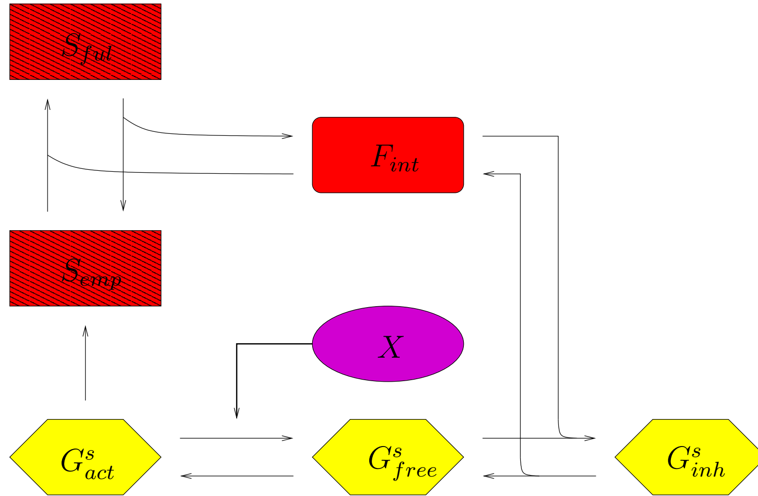


Abbildung 8: a model for the regulation of the iron storage

also have to consider the experiences of the MPI at Magdeburg, where the modelling of the cell's catabolism and diauxic growth has been thoroughly investigated.

Figure 7 shows the uptake of the carbohydrate source C_{ext} mediated by the specific transporter T^c. The iron dependent enzymes F_{cat} are necessary for the energy supply reactions. We denote the energy which is then transferred to the ADP simply by E, because in this context it does not matter what this final energy carrier looks like. The growth of biomass is then an ATP consuming reaction (e.g. polymerisation) of the imported carbohydrate. There is always a decay of a certain rate of DNA, RNA, and protein. We don't distinguish between repair and new synthesis, because from our point of view both processes are nothing but energy and c-source dependent reactions.

The iron storage proteins are another interesting feature we would like to add. The equations should be rather simple. The tricky question will be how one can realize the genetic regulation of the bacterioferritin expression, because it is dependent on Fur and on the growth rate. We are not sure how the cell senses its own growth, but we hope that the biologists can help us with this uncertainty, too. We hope that even in the shake-flask experiment with all its limitations and the usual problem of experiments dealing with

iron concentrations, it is possible to see the effect of the iron storage, if the c-source is given in pulses. One should set up two experiments, one with an E.coli wild type strain and one with a mutant without the ability to produce iron storage proteins. The cell numbers should attain the same level after each sugar pulse, but with different velocities. The next experiment would be to let both strains compete in a single experiment. We expect that the 'fit' wild type strain should displace the weaker mutant.

Figure 8 shows a possible regulation of the expression of the iron storage proteins. It is to be interpreted as a very rough sketch of how we think this subsystem could work. We have not drawn the self-assembly of several bacterioferritins into an iron storing aggregate here. Like in our current system, the transporter genes can be blocked by Fur, but for the expression of the genes a second signal X is required which carries the information that cell growth is in stagnating due to the lack of a carbohydrate source. We do not know yet, if this signal is another repressor which downregulates the activity of these genes under normal growth conditions, or if it is an activator that binds to the DNA in the case of a drop of the c-source concentration. In the image we have drawn the situation where the signal is an activator.

According to [1] (personal contact) there is a partial knowledge about the regulative mechanisms that control the expression of the genes encoding enzymes with a need for iron. We hope that we can extract a better formula for the amount of P_{emp} from this knowledge.

5.2 Qualitative Extension: Other Modelling Concepts

A relatively new approach to model genetic regulation in cells is presented in [4]. It does not follow the chemical modelling tradition but uses an ordinary differential equation with a piecewise linear right hand side. The discontinuities arise from switch functions that either totally enable or disable the gene expression. An easy example for this technique is a negatively autoregulated gene that is switched off if the concentration of its product exceeds a certain threshold. These piecewise linear systems can be regarded as a subclass of the differential inclusions, because at the threshold no single value can be defined for the right hand side. This means that solutions do not exist in the classical but in the so called Filippov sense, and the only interesting dynamics occur inside the threshold planes.

The resulting model is a strong simplification of the system obtained by classical means, and the solutions are even less realistic than what one

obtains by solving the classical system. On the other hand the biological data is usually so bad, that one cannot expect to derive more than qualitative statements. In this case, as for big systems, this method proves very useful, because the numerical integration is easy, and a purely qualitative description of the possible patterns of behaviour of the system can be calculated by a quick algorithm [4].

However, as soon as one does not only look at genetic regulations but at reactions involving two molecules such as conversions of one metabolite into another by an enzyme, this approach is inappropriate. We hope that it is possible to keep some of the advantages of the piecewise linear modelling, if one sets up hybrid models by implementing the regulation of genes as piecewise linear equations, while describing the rest of the cell in the classical way. In artificial cases where the genetic regulation is independent of the metabolome this is clearly true. It would be interesting to investigate what happens in the general case where one does not have this clean separation into genome and metabolome. We give a simple example in the next section.

We are also thinking of using the general differential inclusions to model biological systems. The motivation for this idea is that one normally only has relatively good information about the strong interactions between the reactants. A modeller starts with these and ignores the weak interactions, which might also be important to some degree. We think that it makes sense to model these weak influences by allowing the parameters to fluctuate in a small cuboid around a set of parameters, which seems to be the most likely choice, at every point time. The development of a theory about the numerical solution of such systems is still in progress, but it has been shown [5, 8] that adaptions of some classical solution methods are successful.

5.3 A simple example of a hybrid model

In our model there occurs only one regulation on the genetic level which switches the expression of the transporter on if the iron content of the cell is low and off if it is high respectively. If we treat this expression as a piecewise

linear process, the standard differential equations

$$\begin{aligned}\dot{T}_{free} &= r_5 G_{free} + \ldots \\ \dot{G}_{free} &= -r_4 F_{int} G_{free} + r_{-4} G_{inh} + G_{inh} \frac{\dot{M}}{M} \\ \dot{G}_{inh} &= r_4 F_{int} G_{free} - r_{-4} G_{inh} - G_{inh} \frac{\dot{M}}{M}\end{aligned}$$

simplify to the single equation

$$\dot{T}_{free} = r_5 \omega \, \theta(F_{int}) + \ldots$$

where θ is a switch function given by

$$\theta(x) := \begin{cases} 1 & x < \theta_0, \\ 0 & x \geq \theta_0. \end{cases},$$

ω is the number of genes per unit of biomass, and θ_0 is a threshold value for the internal iron concentration F_{int}. For values of F_{int} higher than θ_0, the expression of transporters is switched off, while for lower values the expression works with maximal speed.

The other reactions of our network are not affected by this change, but we have eliminated equations and parameters which is very useful in big or complicated networks including a lot of genetic regulation processes. The whole system now looks as follows:

$$
\begin{aligned}
\dot{F}_{ext} &= -r_1 F_{ext} T_{free} M \\
\dot{F}_{int} &= r_{-1} T_{occ} - r_{-2} F_{int} T_{free} + r_2 T_{inh} - r_4 F_{int} G_{free} + r_{-4} G_{inh} \\
&\quad -r_3 P_{emp} F_{int} + r_{-3} F_{cat} + d_T (T_{occ} + T_{inh}) - F_{int} \frac{\dot{M}}{M} \\
\dot{F}_{cat} &= r_3 P_{emp} F_{int} - r_{-3} F_{cat} - F_{cat} \frac{\dot{M}}{M} \\
\dot{P}_{emp} &= -r_3 P_{emp} F_{int} + r_{-3} F_{cat} + F_{cat} \frac{\dot{M}}{M} \\
\dot{T}_{free} &= r_5 \omega\, \theta(F_{int}) + r_{-1} T_{occ} - r_1 F_{ext} T_{free} + r_2 T_{inh} - r_{-2} F_{int} T_{free} \\
&\quad -d_T T_{free} - T_{free} \frac{\dot{M}}{M} \\
\dot{T}_{occ} &= r_1 F_{ext} T_{free} - r_{-1} T_{occ} - d_T T_{occ} - T_{occ} \frac{\dot{M}}{M} \\
\dot{T}_{inh} &= r_{-2} F_{int} T_{free} - r_2 T_{inh} - d_T T_{inh} - T_{inh} \frac{\dot{M}}{M} \\
\dot{M} &= \frac{\mu_{max} \left(F_{cat} - \beta_{fe}\right)}{\mu_s + \left(F_{cat} - \beta_{fe}\right)} M
\end{aligned}
$$

Which properties does such a hybrid system have, how much information does the solution contain, and is the quality of this information worth the computational effort in the non-discretized variables? We think that in this context these questions are the right ones to ask, and we will try to answer them in the future.

Literatur

[1] S.C. Andrews, A.K. Robinson, and F. Rodriguez-Quinones. Iron homeostasis in bacteria. *FEMS Microbiol. Rev.*, 27:215–237, 2003.

[2] A. Becker, H. Berges, L. Krol, C. Bruand, S. Rueberg, D. Capela, E. Lauber, E. Meilhoc, F. Ampe, F.J. de Bruijn, J. Fourment, A. Francez-Charlot, D. Kahn, H. Kuester, C. Liebe, A. Puehler, S. Weidner, and J. Batut. Global changes in gene expression in sinorhizobium meliloti 1021 under microoxic and symbiotic conditions. *Mol Plant Microbe Interact*, 17(3):292–303, 2004.

[3] V. Braun and H. Killmann. Bacterial solutions to the iron-supply problem. *TIBS*, 24:104, 1999.

[4] H. de Jong, J.-L. Gouze, C. Hernandez, M. Page, T. Sari, and J. Geiselmann. Dealing with discontinuities in the qualitative simulation of genetic regulatory networks. *Proceedings of Fifteenth European Conference on Artificial Intelligence, ECAI-02, F. van Harmelen (ed.), IOS Press, Amsterdam*, pages 412–416, 2002.

[5] A. Dontchev and F. Lempio. Difference methods for differential inclusions: A survey. *SIAM Review*, 34:263–294, 1992.

[6] O. Ebenhoeh and R. Heinrich. Stoichiometric design of metabolic networks: Multifunctionality, clusters, optimization, weak and strong robustness. *Bull.Math.Biol.*, 65:323–357, 2003.

[7] J.-L. Gouze and T. Sari. A class of piecewise linear differential equations arising in biological models. *Dynamical Systems*, 17:299–316, 2003.

[8] G. Grammel. Towards fully discretized differential inclusions. *Set-Valued Analysis*, 11:1–8, 2003.

[9] K. Hantke. Iron and metal regulation in bacteria. *Current Opinion in Microbiology*, 4:172–177, 2001.

[10] W. Koester. Abc transporter-mediated uptake of iron, siderophores, heme, and vitamin b12. *Res. Microbiol.*, 152:291–301, 2001.

[11] A. Kremling. *Strukturierung zellulaerer Funktionseinheiten – ein signal-orientierter Modellierungsansatz fuer zellulaere Systeme am Beispiel von E.coli*, volume 5 of *Forschungsberichte*. Max-Planck-Institut fuer Dynamik komplexer technischer Systeme, 2002.

[12] A. Kremling, K. Jahreis, J.W. Lengeler, and E.D. Gilles. The organization of metabolic reaction networks: A signal-oriented approach to cellular models. *Metabolic Engineering*, 2:190–200, 2000.

[13] S. Schuster, C. Hilgetag, J.H. Woods, and D.A. Fell. Elementary flux modes in biochemical reaction systems: Algebraic properties, validated calculation procedure and example from nucleotide metabolism. *J. Math. Biol.*, 45:153–181, 2002.

[14] S. Weidner, A. Puehler, and H. Kuester. Genomics insights into symbiontic nitrogen fixation. *Curr Opin Biotechnol*, 14:200–205, 2003.