HUMBOLDT-UNIVERSITÄT ZU BERLIN

MATHEMATISCH-NATURWISSENSCHAFTLICHE FAKULTÄT I

INSTITUT FÜR BIOLOGIE

MASTERARBEIT

ZUM ERWERB DES AKADEMISCHEN GRADES

MASTER OF SCIENCE

Mathematische Modellierung logischer Funktionen in genetisch manipulierten *S. cerevisiae*

Mathematical modeling of synthetic logic gates in *S. cerevisiae*

vorgelegt von

Adriana SUPADY

angefertigt in der Arbeitsgruppe THEORETISCHE BIOPHYSIK

am Institut für Biologie

Berlin, im März 2012

Contents

Abstract

One of the major goals of synthetic biology is the construction of programmable biocomputing devices. An example of a successful implementation is a recently published system of synthetic logic gates employing a library of engineered yeast cells. In this work we present a quantitative analysis of this system and show how mathematical modeling can support its further development. Deterministic and stochastic approaches are applied to model logic gates implementing the following Boolean functions: IDENTITY, NOT, OR and IMPLIES. A sensitivity analysis is performed in order to identify the most significant parameters for the system output. Furthermore, we investigate the noise distribution in the system and the ability of our models to reproduce the experimental data. We also show how the output of the system can be optimized by e.g. changing the density of the cell culture. Finally, we propose an extended model that implements a function following the three-value logic .

Zusammenfassung

Eines der Ziele der synthetischen Biologie ist die Konstruktion programmierbarer Elemente, die aus biologischen Einheiten aufgebaut sind. Ein Beispiel für die erfolgreiche Umsetzung dieses Konzeptes ist die vor kurzem veröffentlichte Implementierung logischer Funktionen in genetisch modifizierten Hefezellen. In der vorliegenden Arbeit präsentieren wir eine quantitative Analyse dieses Systems und zeigen, wie mathematische Modellierung dessen Weiterentwicklung unterstützen kann. Wir stellen deterministische und stochastische Modelle vor, die die Funktionen IDENTITÄT, NEGATION, DISJUNKTION und IMPLIKATION implementieren. Mit Hilfe einer Sensitivitätsanalyse werden für das System kritische Parameter identifiziert. Darüber hinaus untersuchen wir die Verteilung des Rauschens im System und überprüfen, ob die Modelle die experimentellen Daten reproduzieren können. Zudem zeigen wir, wie das System durch die Änderung der Zelldichte optimiert werden kann. Letztendlich schlagen wir ein erweitertes Modell zur Implementierung einer dreiwertigen logischen Funktion in zellulären Systemen vor.

Chapter 1

Introduction

Contents

The design and creation of a computational tool functioning like an electronic device build from biological units is a joint effort of systems and synthetic biology. The work of Regot and colleagues [1] is an excellent example of how this two distinct fields meet. Biologically engineered cells were combined together into complex systems in order to be able to perform simple logic operations. The aim of this work is to provide mathematical models of these artificial constructs, test their behavior, make predictions and suggest modifications.

1.1 Synthetic biology

Description Synthetic biology is a young field that combines molecular biology and engineering in order to design artificial biological systems. It provides a great opportunity to test the present understanding of natural biological mechanisms [2, 3, 4]. The first step in the engineering process is the choice of a biological function of interest and designing a system that is able to perform this function. As next, the compounds that can be used for the construction have to be found. Then the system can be established and evaluated in order to prove if the function is executed properly [5]. The goal of the synthetic biological engineering is to construct customized biological devices that are able to recognize the input, process the information, perform the desired function and produce an output [6]. Biological devices that communicate with each other and are placed in an appropriate cellular context can build biological modules and perform complex tasks. One of the first and best know examples of a synthetic system that performs a particular function is the repressilator that was built in *E. coli* from three transcriptional repressors systems [7]. Due to the negative-feedback loop, the network oscillates and activates the synthesis of the GFP that is the desired system function. The design and the implementation of the repressilator that performs a designated function is considered to be a milestone of synthetic biology.

Limitations Although one of the goals of synthetic biology is a precise control in biological systems there are two main difficulties that arise while engineering biological and not electronic systems. As first, the biological mechanisms that perform desirable functions are rarely know in

details. Secondly, in any biological system the environment has a great impact on its functionality and changes in the outer world can lead to the loss of the function. Despite these difficulties many researchers succeeded in manipulation and extension of the existing signaling pathways in the cells. For instance, it was shown that synthetically constructed feedback loops can be applied to reprogram the cellular responses [8].

Applications Without doubt synthetic biology has a great potential as a field. New compounds and networks can be established and used for rewiring and reprogramming of organisms [9]. The applications range from therapeutics through biosensing to the production of biomaterials and biofuels. One possible application could be engineered bacteria that are able to recognize tumors and then invade the cancerous cells [10]. Another example is the fact that synthetic pathways could potentially be used for microbial production of an antimalarial drug precursor [11].

1.2 Logic gates

Logic gates are devices that implement Boolean functions converting logic inputs into a single logic output. They employ binary logic where only two values are possible: (i) 0-false (no signal or low signal) and (ii) 1-true (high signal) [12]. Although the concept mainly refers to digital systems in electronics it is also possible to build biological logic gates from e.g. gene circuits [13]. Biological logic gates are far more complicated to construct than the electronic circuits. The main difficulty arises from the need of wiring the biological units that perform certain tasks. Constructing of logic gates within a single cell is coupled with diverse limitations, like e.g.: (i) crosstalk between the pathways as well as (ii) uncontrolled noise propagation. These difficulties can be evaded by introducing cellular compartmentalization that is shown in **Figure 1.1**. This novel strategy was proposed and implemented in engineered *Saccharomyces cerevisiae* strains by Regot and colleagues [1].

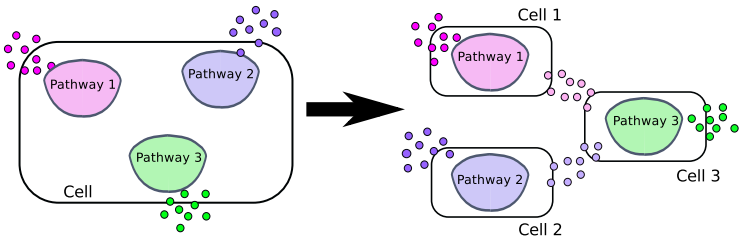


Figure 1.1: Cellular compartmentalization. In this novel approach each cell type performs exactly one logic operation. The cells are wired by the diffusible output molecules.

As shown, each cell in the system performs only one function and responds only to a particular input. This solves the problem of wiring between the cells and the logic gate can perform desired function with a high fidelity of the information flow.

Tamsir and colleagues [14] also used a multicellular strategy to design their logic gate in *E. coli*. They compartmentalized the logic gate into separate *E. coli* strains and employed quorum signaling to enable the cell-to-cell communication.

1.3 *Saccharomyces cerevisiae* as a model system

The yeast *Saccharomyces cerevisiae* is one of the best studied eukaryotic model organisms. Many human proteins were first discovered and studied as their homologs in yeast, e.g. diverse signaling proteins [15]. *S. cerevisiae* can be cultured at low cost, has a short doubling time and can be easily manipulated (by adding new genes or creating gene knockouts). These features make *S. cerevisiae* attractive for studying of new functions.
In order to sense and to respond to the changes in the environment as well as to integrate the external and internal signals yeast cells use signal transduction pathways [16]. The precise functioning of these pathways is crucial for the survival of the cells. *S.cerevisiae* often employs mitogen activated protein kinases (MAPKs) to transmit the information about external stimuli and to cause intracellular responses [17]. To give an example, MAPK cascade can act as a link between the upstream signaling (e.g. receptor, G-proteins) and downstream effectors (e.g. regulators of gene expression). The MAPKs usually build a three-step-cascade: MAPKKK receives the signal and can phosphorylate the MAPKK which can in turn phosphorylate the MAPK. The latter usually acts as a final trigger for diverse responses.
In the following subsections two signal transduction systems of yeast that employ MAPKs will be described: (i) the pheromone pathway, responsible for mating of the yeast cells and (ii) the high-osmolarity glycerol (HOG) pathway, activated in response to osmotic stress. Both are crucial for understanding of the model system that will be presented in the latter chapters of this work.

1.3.1 Pheromone pathway

The pheromone pathway in *S. cerevisiae* is one of the best studied pathways [18, 19, 20]. The budding yeast cells may exist in two pheromone producing haploid forms: MATa cells that produce a-factor and MATα cells that produce α-factor. MATa cells carry Ste2 receptor and are able to sense the α-factor. Similarly, MATα cells carry Ste3 receptor, to which the a-factor can bind in order to stimulate the mating process. The active receptor can interact with the Gα subunit of the the heterotrimeric G-protein (consisting of Gα, Gβ and Gγ subunits) and activate the G-protein cycle that participates in the transmission of the signal. During this process the heterodimer G$\beta\gamma$ is released and can bind and activate other components of the pathway e.g the scaffold protein Ste5 that recruits the elements of the MAPK cascade to the plasma membrane. The activation of the signaling cascade leads to the phosphorylation of the MAPK Fus3 that shuttles between the cytoplasm and nucleus. The active Fus3 is responsible for phosphorylation of diverse proteins that enable mating of the cells e.g. Far1 (responsible for the arrest of the cell cycle) or Ste12 (participates in the expression of pheromone-induced genes)[21].

1.3.2 HOG pathway

Yeast cells live in a permanently changing environment. In order to survive special adaptation mechanisms must have evolved. One of them is the ability of the cells to respond to the changes in the osmotic conditions. This response employs the high-osmolarity glycerol (HOG) pathway [22, 23]. In presence of osmotic shock, two osmosensors Sln1 and Sho1 become active and phosphorylate then the MAPKK Pbs2 which in turn phopshoprylates the MAPK Hog1 that

shuttles between the cytoplasm and nucleus. The nuclear active Hog1 controls the expression of over 600 genes via phosphorylation of osmoresponsive transcription factors e.g. Hot1 [24]. The active cytoplasmic Hog1 contributes to the accumulation of internal glycerol that compensates the external osmolarity, regulates the cell volume and the turgor pressure that prevents the dehydration of the cell and ensures osmotic stabilization.

Chapter 2

Methods

Contents

2.1 Modeling of biological processes

Mathematical models gives us the opportunity to understand the intrinsic nature and dynamics of biological processes. Mathematical modeling enables us also to make predictions as well as to integrate the experimental knowledge into a common framework [16, 25]. A model should represent a biological process with appropriate precision in order to explain desired features or functions. Models can vary in art and complexity depending on the purpose of modeling [26]. In other words, as first a hypothesis that we want to test has to be formulated and then, after choosing an appropriate modeling strategy, the model can be established and validated. In the following subsections we describe two modeling approaches that we use in our work: kinetic modeling with ordinary differential equations (ODEs) and stochastic modeling.

2.1.1 Deterministic modeling

One of the most common techniques to model the temporal changes in the state variables is to describe the dynamics of the system with a set of ODEs [25].
The temporal changes in the system of n components with concentrations c and l distinct kinetic parameters p can be described as follows:

$$\frac{dc_i}{dt} = \dot{c_i} = f_i(c_1, ..., c_n, p_1, ..., p_l, t) \qquad i = 1, ..., n$$

The concentrations of the species c_i are time-dependent functions of concentrations of other chemical species and model parameters. The initial concentrations of the model compounds and kinetic parameters rarely can be all obtained from the literature. The missing values have to be fitted (estimated basing on available experimental data). The aim of parameter estimation is to find such a set of parameters so that the simulations of the model match the experimental data set with highest possible accuracy. The missing parameter values of the ODE models

presented in this work were obtained with the tool COPASI [27] and the method Evolutionary Programming [28].
All details to reactions, parameters and initial concentrations that we use in our work will be presented in the chapter **Model construction**.

2.1.2 Stochastic modeling

In the deterministic approach we assume the compound quantities to be continuous values. It is a good approximation for compounds with high abundances. Nevertheless, some biochemical species exist in a small copy number. Therefore, there is a need for a different approach that could model species quantities as discrete values. Due to a small number of molecules significant stochastic effects can occur [29]. One of the best studied and optimized methods for stochastic modeling is the stochastic simulation algorithm (SSA) employing the direct method [30] that answers the following questions: (i) which reaction occurs as next and (ii) when does it occur. The SSA is a numerical method that allows to simulate well-stirred reacting systems in exact accordance with the chemical master equation (CME) [31].
A chemical reaction is defined as a random collision of two molecules in a collision volume. Having the initial amount of molecules of several biochemical species and reaction probability constants it is possible to simulate the dynamics of the system, reaction by reaction. The algorithm includes following steps:

1. Set time=0. Initialize the abundance of the molecules and reaction propensities.
2. Generate random numbers in order to identify the next reaction to fire and the time interval.
3. Update the time by adding the length of the generated time interval. Update the abundances of molecules.
4. Continue till the simulation time has been exceeded.

Each run of the simulation is different and is not necessarily reflecting the behavior of the whole system. In order to answer the question, how the system performs, many simulations have to be run and averaged. This also gives the possibility to calculate the distribution of the results in order to quantify the noise in the system that is an important feature of biological systems. For some biological questions it is of more importance to know how the results are distributed than to know the average of them. All stochastic simulations in this work were performed with the tool CAIN [32].

2.1.3 Comparison of the approaches

Both methods described in previous subsections are relatively fast and easy in implementation. There are many tools available that can be used to implement both deterministic and stochastic models. The deterministic modeling with ODEs is also very fast in computing and easy to analyze. It is most suitable for systems with high molecule numbers, where the fluctuations are insignificant. We must remember that the result of an ODE model simulation is always an average for the whole system.

The stochastic approach has a lot of advantages but also limitations [30]. The SSA is an exact method allowing all fluctuations to occur. In contrast to ODE modeling, SSA does not approximate the time steps and therefore allows sudden and sharp changes of the system. Only the SSA gives the possibility to simulate the behavior of a single molecule. However, simulating each molecular reaction is often too time-consuming for simulations of more complex systems. Summing up, the choice of the method depends on the problem to solve and its complexity. In the following chapters we employ both methods to test our hypotheses. We construct and simulate ODE models, a stochastic model and a combined hybrid model.

2.2 Boolean logic

As already mentioned in the **Introduction**, logic gates perform functions following Boolean logic that uses only two values: 0 (false) and 1 (true). The number of functions B that can be defined with the Boolean algebra depends on the number of input variables n and is given with following equation:

$$B(n) = 2^{2^n}$$

For a single input (p) there are four different boolean functions existing. The output values of the functions with corresponding descriptions are listed in **Table 2.1**.

Input	p	0	1	Description
Output	FALSE	0	0	Output is always false
	IDENTITY	0	1	Output is true only when p is true
	NOT	1	0	Output is false only when p is true
	TRUE	1	1	Output is always true

Table 2.1: Logic operations for a single input(p). Four different functions can be defined for a single input. They are listed with the corresponding output values and descriptions.

If an input consist of two boolean variables (p and q), 16 possible boolean functions exist. Six of them are listed with descriptions in **Table 2.2**.

Input	p	0	1	0	1	Description
	q	0	0	1	1	
Output	AND	0	0	0	1	Output is true only when both p and q are true
	NAND	1	1	1	0	Output is false only when both p and q are true
	OR	0	1	1	1	Output is true when at least one of p and q is true
	NOR	1	0	0	0	Output is true only if both p and q are false
	IMPLIES	1	0	1	1	Output is false only if p is true and q false
	N-IMPLIES	0	1	0	0	Output is true only if p is true and q false

Table 2.2: Logic operations for an input operating with two values (p and q). Six of 16 different functions that can be defined for an input with two values are listed with the corresponding output values and descriptions.

We apply Boolean functions to evaluate the performance of the logic gates and to interpret the results.

Chapter 3

Model construction

Contents

3.1 Deterministic models of single cells

Mathematical models of four different cells have been established and afterwards combined into four gates able to express following logic functions: IDENTITY, NOT, OR and IMPLIES. The cells are constructed in a way that each cell can act as either a sender or reporter cells. The structures of the models of the cells are presented in **Figure 3.1**.

The sender cells are able to respond to external stimuli (salt, galactose or doxycycline) and to produce the α-factor. The reporter cell is able to sense the α-factor in the medium and than produce GFP. As first the deterministic models of single cells were created based on the parameters found in the literature. The remaining parameters were fitted to diverse literature data. The following subsections describe in detail how the cells were engineered [1] and how are they modeled and show also their behavior when stimulated with characteristic input concentration values. Tables with kinetic equations, parameter values and initial concentration values with all references are collected in the **Appendix**. The references to the data sets that were used for parameter estimation are also collected in the **Appendix**. All models were implemented and simulated in COPASI [27].

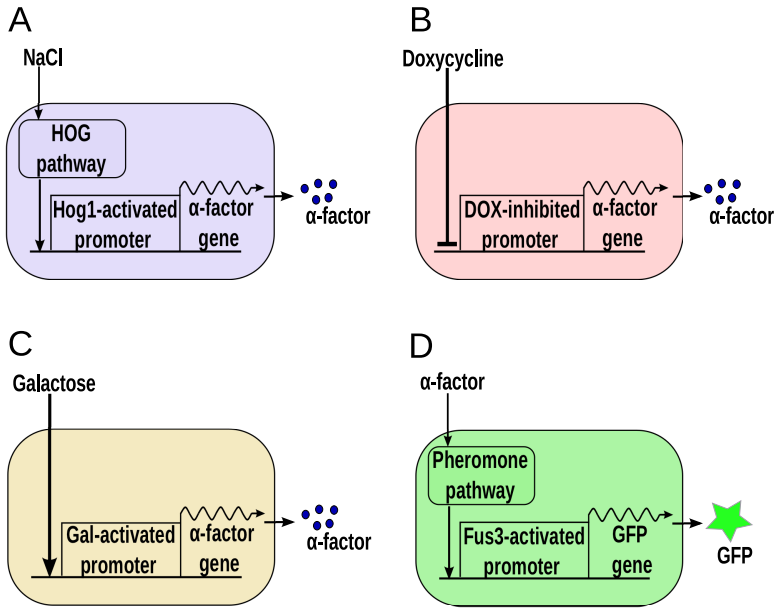


Figure 3.1: Schemes of the structures of the sender cells: (**A**) salt-cell, (**B**) dox-cell, (**C**) gal-cell and the (**D**) reporter cell.

3.1.1 Salt-cell

The salt-cell is a MATα cell with deleted Ste3 receptor. This deletion prevents mating with MATa cells and the production of endogenous α-factor. However, the salt-cell can produce α-factor in the presence of NaCl because the *MFα1* gene is under the control of the osmostress-induced *STL1* promoter. Additionally, the cell carries a *fps1-Δ1* mutation which prevents fast glycerol accumulation in the cell and causes an oversensitivity to the osmostress [33].

Model The structure of the model is shown in **Figure 3.2**.

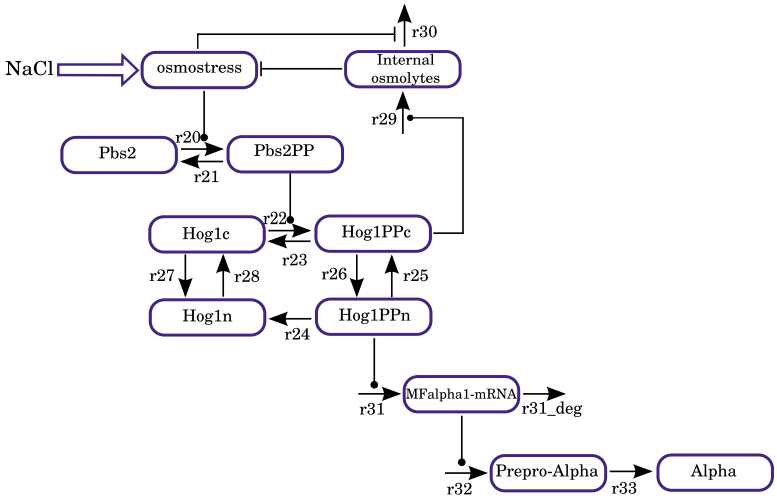


Figure 3.2: Salt-cell: wiring in the system. Each arrow represents a chemical reaction. Lines with dots at the end are modulation reactions. Lines with perpendicular lines at the end are negative regulations. Biochemical species with n at the end of the name are nuclear species. All other components are in the cytoplasm. The system is activated in the presence of salt and is able of producing α-factor.

The osmostress is build upon the presence of NaCl in the medium and can be compensated by the accumulation of internal osmolytes. In the presence of osmostress the MAPKK Pbs2 becomes phosphorylated (**r20**). The phosphorylated form Pbs2PP can (i) be dephosphorylated (**r21**) or (ii) can activate, via phosphorylation, the cytoplasmic MAPK Hog1 (**r22**) which can also be dephosporylated (**r23**). It is not possible to phosporylate Hog1 in the nucleus thus only the reversible reaction can occur (**r24**). All four forms of Hog1 can shuttle between cytoplasm and nucleus (**r25-r28**). The presence of active cytoplasmic Hog1 enables the accumulation of internal osmolytes (**r29**). The internal osmolytes can also release the cell (**r30**) if there is no osmostress present. The nuclear active Hog1 activates the *STL1* promoter which induces the production of *MFα1*-mRNA (**r31**). The *MFα1*-mRNA induces the production of the polypeptides needed for the synthesis of the α-factor - PreproAlpha (**r32**) which can then be transformed into Alpha (**r33**). The *MFα1*-mRNA is being constantly degraded (**r31deg**).

The produced Alpha is being diluted when exported from the cytoplasm to the medium. Therefore we introduce a parameter *dilution*. The value for this parameter is calculated from the values of the volumes of the compartments and the cell culture density. The initial value of the cell culture density in our models is set to $5 \cdot 10^6 \frac{cells}{ml}$ that correspond to a dilution factor of 13800. However, the density of the cell culture changes during the time due to the cell division. We assume that the cells have a doubling time of four hours. As a consequence, the density of the culture increases during the simulation that results in a decrease in the value of the dilution factor. We model the value of the dilution factor with following equation:

$$dilution = 13800 \cdot e^{-0.173 \cdot Time[h]}$$

The concentration of the produced α-factor (also called Alpha in culture) is modeled with following equation:

$$[\alpha\text{-factor}] = \frac{[Alpha]}{dilution}$$

Deterministic simulations We run deterministic simulations of the ODE model of the salt-cell for four hours. We payed particular attention to following processes: the activation of nuclear Hog1 and the production of the α-factor under diverse osmotic conditions (0.0, 0.05, 0.1, 0.2, 0.4 and 0.6 M NaCl).

Figure 3.3 presents the activation of the nuclear Hog1 (Hog1PPn) that reaches its maximum within first minutes of stimulation. The value of the reached maximum is concentration dependent and increases with the concentration of salt. The increase in the osmotic pressure results also in a delay in reaching of the the steady-state because the system needs more time for adaptation.

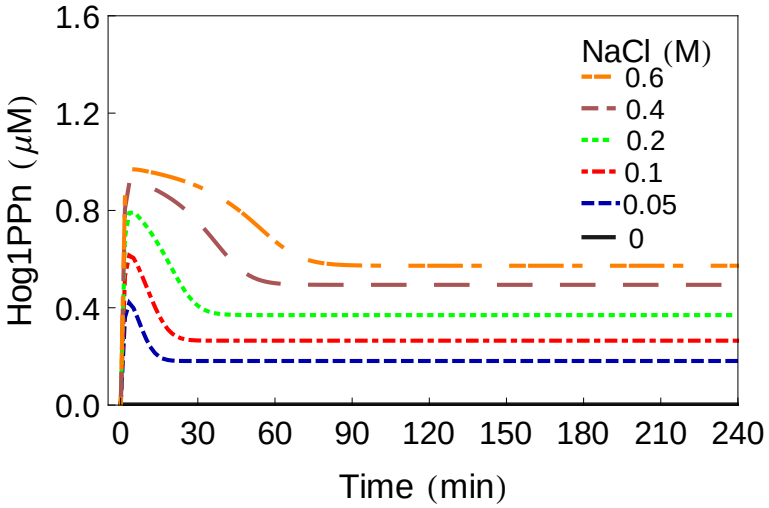


Figure 3.3: Simulation of the deterministic salt-cell model. The concentration of Hog1PPn under stimulation with different salt concentrations is plotted over four hours. The higher the osmotic stress the higher peak and plateau are reached.

Figure 3.4 illustrates the differences in the production of α-factor by the salt-cell model for different salt concentrations. The amount of produced α-factor after four hours of simulation ranges from 0.7 nM (for 0.05 M NaCl) to 2.5 nM (for 0.6 M NaCl).

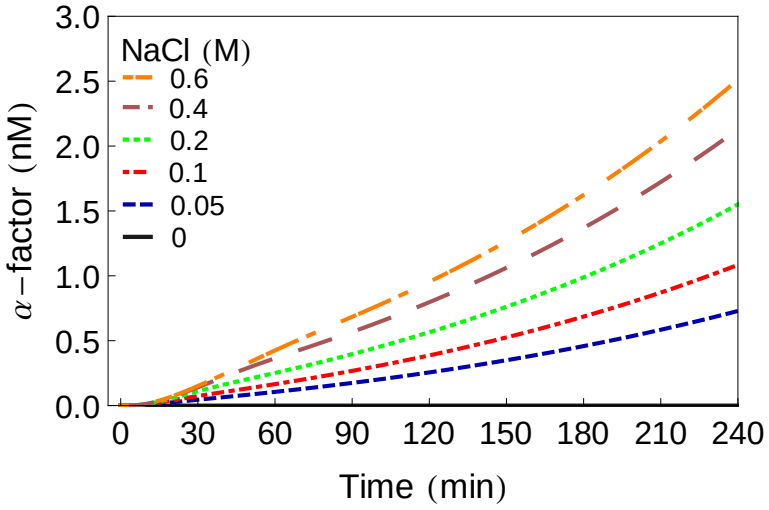


Figure 3.4: Simulation of the deterministic salt-cell model. The concentration of α-factor under stimulation with different salt concentrations is plotted over four hours. The higher the osmotic stress the more α-factor is produced.

3.1.2 Dox-cell

The dox-cell is a MATα cell with deleted Ste3 receptor that prevents mating with MATa cells and the production of endogenous α-factor. The *MFα1* gene is under the control of two TetOperators that enable a continuous α-factor expression in the absence of doxycycline. The presence of appropriate doxycycline concentration represses very quickly the production of α-factor .

Model The structure of the dox-cell model is shown in **Figure 3.5**.

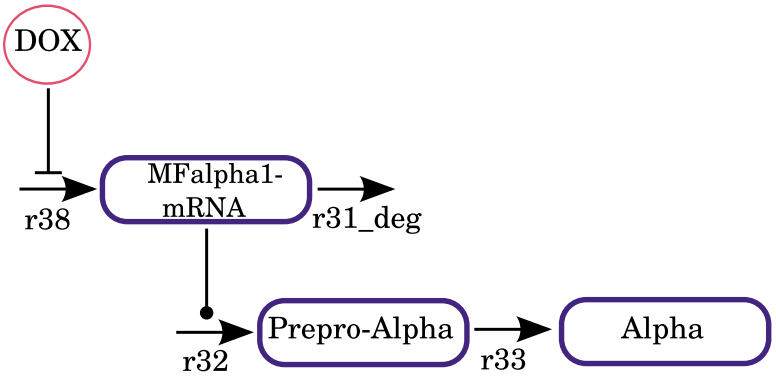


Figure 3.5: Dox-cell: wiring in the system. Each arrow represents a chemical reaction. Lines with dots at the end are modulation reactions. Lines with perpendicular lines at the end are negative regulations. The system is able of producing α-factor. This process is negatively regulated by doxycycline.

Doxycycline enters the nucleus and represses the expression of *MFα1* gene **(r38)**. The reactions contributing to the production of Alpha (**r31deg**, **r32**, **r33**) are the same for all sender cells and were already described in previous subsection about the salt-cell.

Deterministic simulations We run deterministic simulation of the ODE model of the dox-cell and simulated the behavior of the cell upon diverse doxycycline concentrations (0.0, 0.05, 0.1, 0.5, 1.0, 5.0 and 10.0 μg/ml). **Figure 3.6** presents the amount of produced α-factor by the dox-cell stimulated with different doxycycline concentrations.

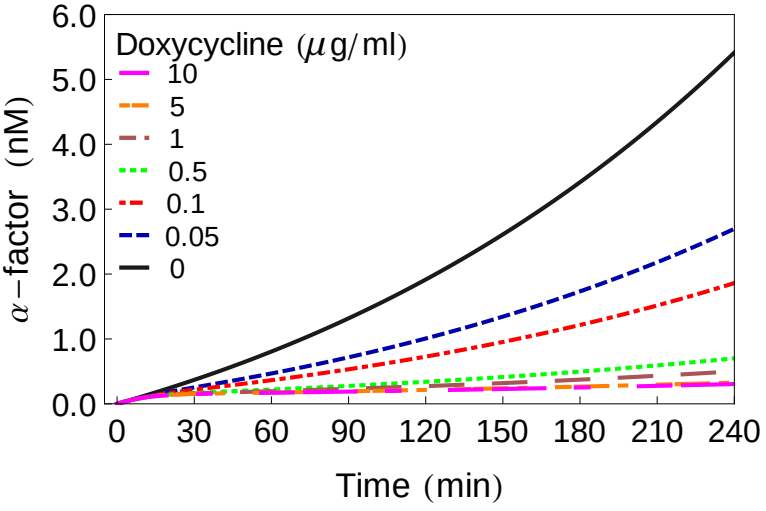


Figure 3.6: Simulation of the deterministic dox-cell model. The concentration of α-factor under stimulation with different doxycycline concentrations is plotted over four hours. The higher the concentration of the doxycycline the less α-factor is produced.

The simulation reveals that already a low concentration of doxycycline (0.05 μg/ml) significantly decreases the amount of produced α-factor produced by the dox-cell.

3.1.3 Gal-cell

The gal-cell is a MATα cell with deleted Ste3 receptor that prevents mating with MATa cells and the production of endogenous α-factor. The *MFα1* gene is placed under the control of *GAL1* promoter and enables the expression of α-factor in the presence of galactose. Addition of glucose represses the production.

Model The structure of the model is shown in **Figure 3.7**.

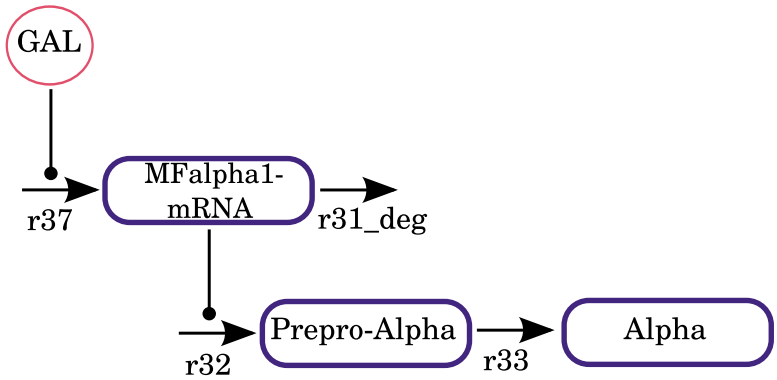


Figure 3.7: Gal-cell: wiring in the system. Each arrow represents a chemical reaction. Lines with dots at the end are modulation reactions. The system is able of producing α-factor. This process is stimulated in the presence of galactose

Galactose activates the promoter and enables the production of *MFα1*-mRNA (**r37**). The reactions contributing to the production of α-factor (**r31deg**, **r32**, **r33**) are the same for all sender cells and were already described in the subsection about salt-cell.

Deterministic simulations We simulated the behavior of the gal-cell performing deterministic simulations of the production of α-factor for different galactose concentrations (0.0, 0.5, 1.0, 2.0, 4.0 %) and present the result in **Figure 3.8**

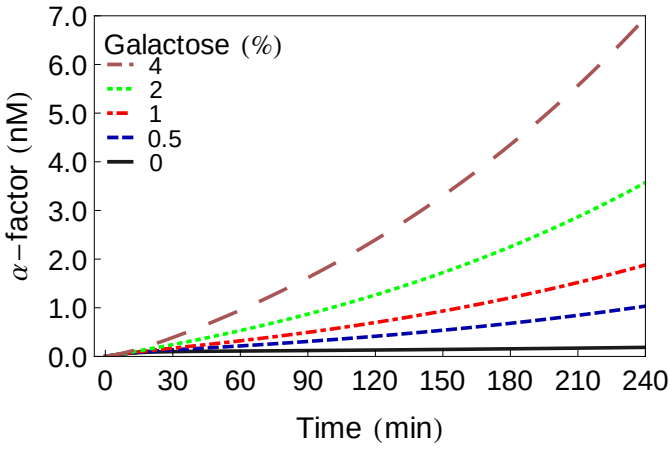


Figure 3.8: Simulation of the deterministic gal-cell model. The concentration of α-factor under stimulation with different galactose concentrations is plotted over four hours. The higher the concentration of the doxycycline the more α-factor is produced.

The rate of accumulation of the α-factor increases with increasing galactose concentration.

3.1.4 Reporter cell

The reporter cell is a MATa cell that carries a *bar1*Δ mutation. BAR1 is an enzyme responsible for the cleavage of α-factor in the extracellular space of MATa cells and its deletion leads to an increased sensitivity to the α-factor [34]. The *GFP* gene is placed under the control of *FUS1* promoter that is induced during the pheromone stimulation. Therefore the reporter cell is able to produce GFP in the presence of pheromone.

Model The structure of the model is shown in **Figure 3.9**.

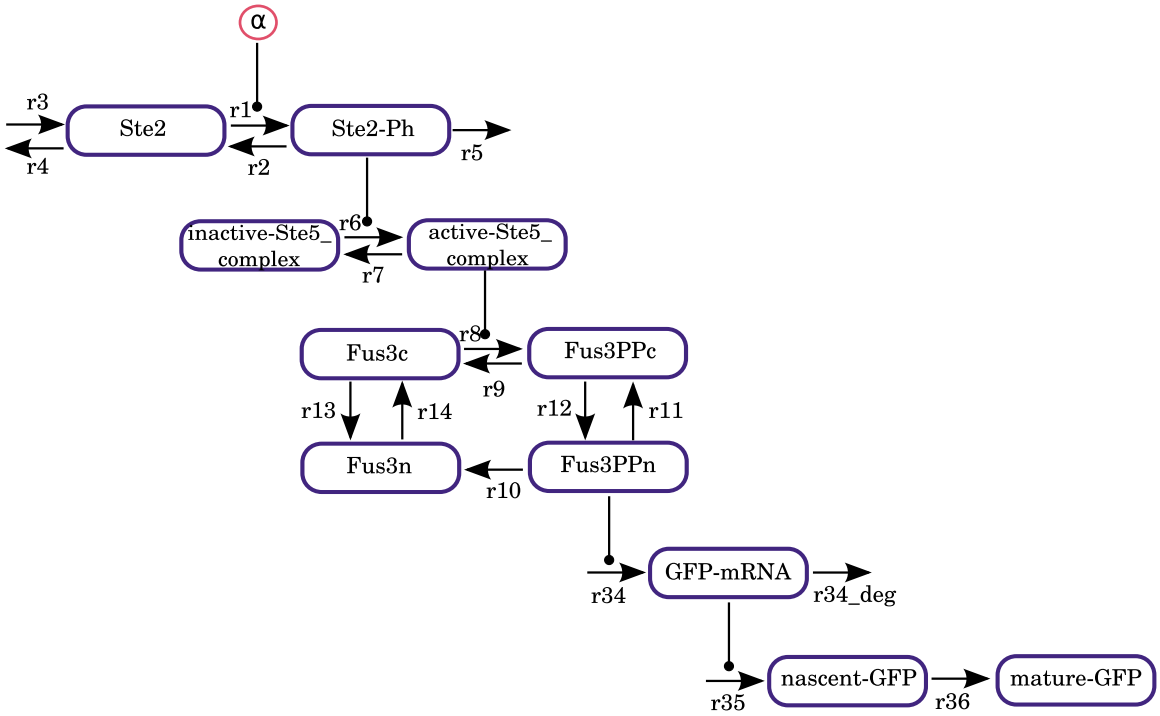


Figure 3.9: Reporter-cell: wiring in the system. Each arrow represents a chemical reaction. Lines with dots at the end are modulation reactions. Biochemical species with n at the end of the name are nuclear species. All other components are in the cytoplasm. The system is able of producing GFP when stimulated with α-factor.

The pathway is activated in the presence of α-factor that binds to the Ste2 receptor (**r1**). There is an internal production of the receptor included (**r3**) as well as degradation of both forms of Ste2 can occur (**r4, r5**). The active form of Ste2 can be deactivated (**r2**) or can switch on the MAPK cascade. For simplification, the MAPK cascade is modeled as a single step: the complex can be activated (**r6**) and deactivated (**r7**). The active complex can phosphorylate the cytoplasmic MAPK Fus3. We can distinguish four forms of Fus3 (cytoplasmic active and inactive as well as nuclear active and inactive) that can shuttle between cytoplasm and nucleus (**r11-r14**). Both cytoplasmic and nuclear forms of Fus3 can be dephosphorylated (**r9, r10**). The nuclear active Fus3 activates the *FUS1* promoter which induces the production of *GFP*-mRNA (**r34**). The *GFP*-mRNA induces the production of nascent GFP (**r35**) which can then be transformed into the mature GFP (**r36**). The *GFP*-mRNA is being constantly degraded (**r34deg**).

Deterministic simulations To investigate further the reporter cell model several simulations have been performed. **Figure 3.10** shows the dynamics of the activation of nuclear Fus3 (Fus3PPn) after stimulation with different α-factor (pheromone) concentrations (0.01 nM, 0.1nM, 1 nM, 10 nM and 100 nM).

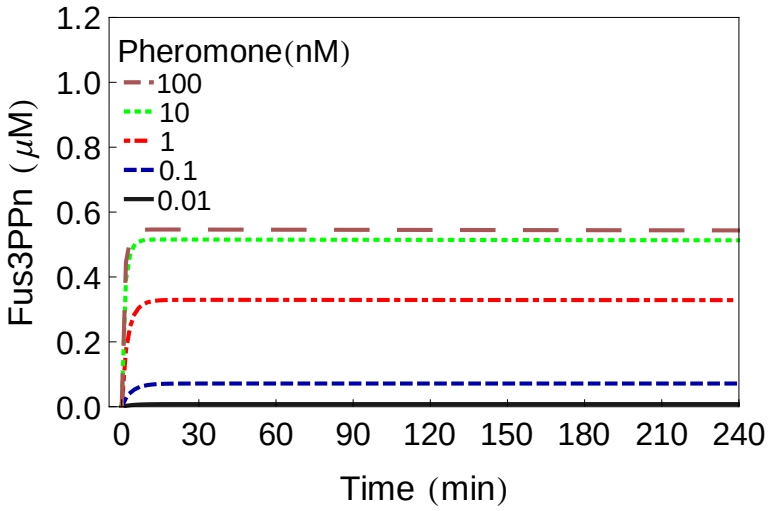


Figure 3.10: Simulation of the deterministic reporter cell model. The concentration of Fus3PPn under stimulation with different α-factor concentrations is plotted over four hours. The system reaches the steady-state within 15 minutes. Its value depends on the α-factor concentration.

In all cases the model reaches the steady state within the first 15 minutes of simulation. The value of the steady state concentration of activated nuclear Fus3 increases with increasing α-factor concentration. Finally, **Figure 3.11** presents the production of mature GFP for different pheromone concentrations.

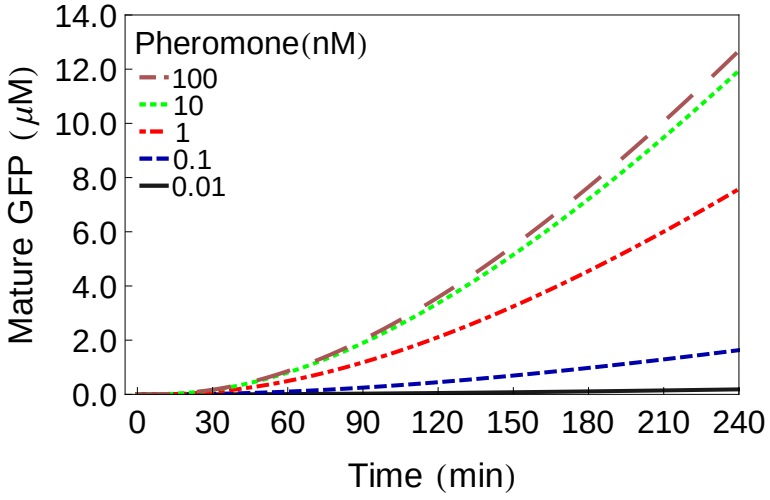


Figure 3.11: Simulation of the deterministic reporter cell model. The concentration of mature GFP under stimulation with different α-factor concentrations is plotted over four hours. The higher the concentration of the α-factor the more GFP is produced.

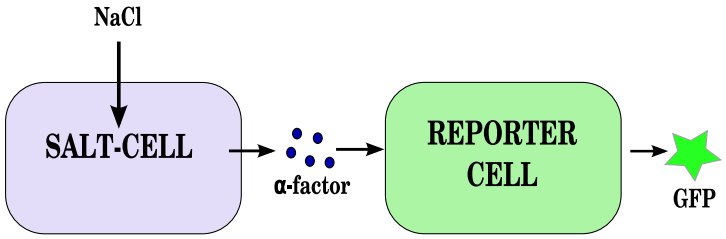
The amount of accumulated mature GFP increases with the increasing pheromone concentration.

3.2 Construction of gates

The in previous section described and modeled single cells can be combined to build gates that perform logic functions. In the following subsections each of the four types of gates that we investigated will be described in detail.

3.2.1 IDENTITY gate

IDENTITY gate is a 2-cell gate built from the salt-cell and reporter cell. **Figure 3.12** presents the scheme of the gate and the corresponding truth table (a truth table is used for computing of the values of logical expression for all possible combinations of arguments; it is composed from one column for each variable and one column for the resulting value).



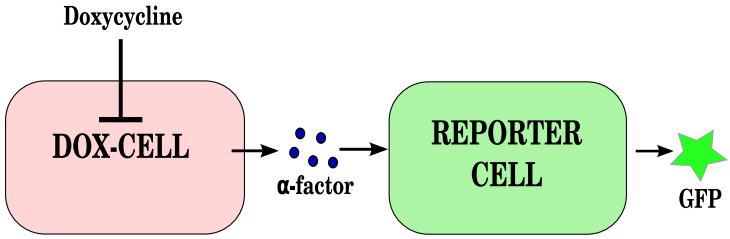
Input	Output
1	1
0	0

Figure 3.12: IDENTITY Gate: Structure and truth table. The gate produces GFP only in the presence of NaCl.

The IDENTITY gate produces output (GFP) only in the presence of input (NaCl). Alternatively, an IDENTITY gate can be also build from the gal-cell and the reporter cell.

3.2.2 NOT gate

The NOT gate is built from a dox-cell and a reporter cell. The system is presented in **Figure 3.13**.



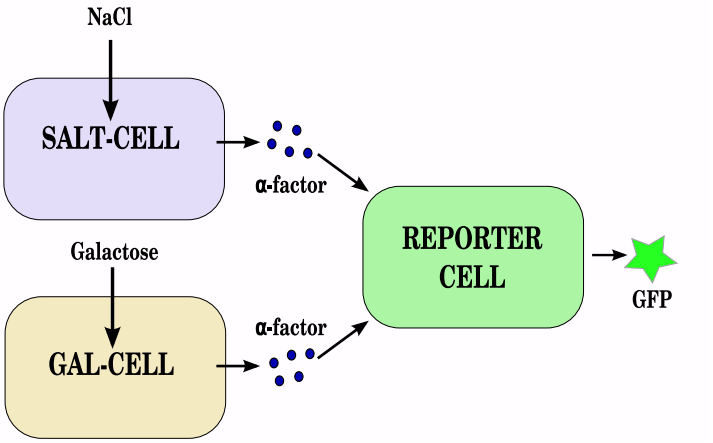
Input	Output
1	0
0	1

Figure 3.13: NOT Gate: Structure and truth table. The gate produces GFP only in the absence of doxycycline.

The NOT gate is constantly active in the absence of doxycycline. Once there is doxycycline in the medium, there will be no output produced.

3.2.3 OR gate

The OR gate is a 3-cell gate built from two sender cells: salt-cell and gal-cell and the reporter cell. The **Figure 3.14** illustrates the system.



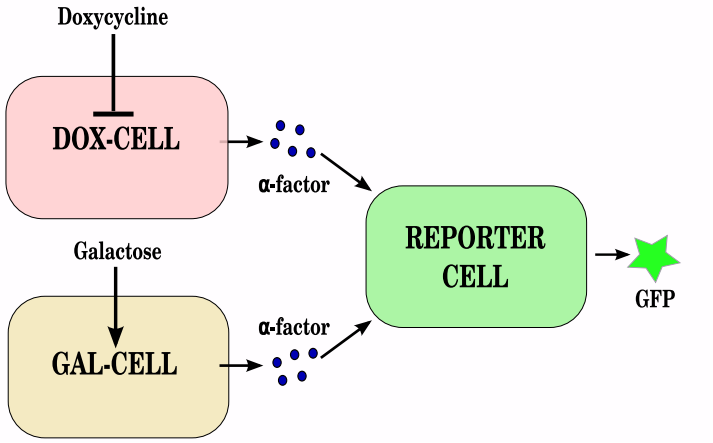
Input 1	Input 2	Output
1	1	1
1	0	1
0	1	1
0	0	0

Figure 3.14: OR Gate: Structure and truth table. The gate does not produces GFP only in the absence of both salt and galactose.

If the salt-cell is stimulated with salt, it will produce α-factor that will lead to GFP production by the reporter cell. Similar holds for the gal-cell: if stimulated with galactose, it will produce α-factor and, subsequently, induce GFP production in the reporter cell. Only if neither salt-cell nor gal-cell is stimulated there will be no GFP been produced. Therefore, the system expresses the OR function.

3.2.4 IMPLIES gate

The IMPLIES gate is built from gal-cell, dox-cell and reporter cell. The gate mechanism is presented in **Figure 3.15** .



Input 1	Input 2	Output
1	1	1
1	0	0
0	1	1
0	0	1

Figure 3.15: IMPLIES Gate: Structure and truth table. The gate does not produce output only in the presence of doxycycline and simultaneous absence of galactose.

In the presence of galactose in the medium, the gal-cell is stimulated and produces α-factor that leads to GFP production by the reporter cell. In the absence of galactose, there are two possible scenarios: (i) doxycycline is absent and the dox-cell produces continuously the α-factor and the reporter cell can produce the GFP or (ii) doxycycline is present in the medium and represses the α-factor production and prevents the system from producing output. Therefore, the system follows the IMPLIES logic.

3.3 Stochastic model of the reporter cell

3.3.1 Full model

The only available experimental data for the synthetically engineered gates are the population data from fluorescence-activated cell sorting (FACS) analysis counting cells that produce detectable amount of GFP [1]. In order to mimic the behavior of the cell population we decided to model the reporter cell in a stochastic manner. This approach is justified by the fact that some of the chemical species in the reporter cell have low abundance and the fluctuations on the single molecule level cannot be overlooked.

Model The structure of the stochastic reporter cell model is the same as the structure of the corresponding deterministic model (**Figure 3.9**). The system was implemented with the help of the tool CAIN [32] and simulated with the build-in function "direct method". Parameters were adopted or appropriately recalculated from the deterministic model and are listed in **Appendix**.

Stochastic simulation In order to investigate the behavior of the constructed model we performed stochastic simulations. Due to the very long computation time we decided to apply quasi-steady-state assumption (QSSA) [35] for the Ste2 receptor to shorten the execution time. The QSSA can be applied here because two prerequisites are fulfilled. As first, the binding of the α-factor to the receptor follows a very fast dynamics resulting in a rapid reaching of a steady state. Secondly, Ste2 is present in a large molecule number in the cell and the fluctuations at the single cell level can be neglected. After applying the QSSA, we simulated 100 trajectories for the reporter cell stimulated with constant α-factor concentrations (0.5, 2.5, 5 nM). As an example, how a result of a stochastic simulation looks like, we present the trajectories for two species (Fus3PPn and *GFP*mRNA) of the reporter cell in **Figure 3.16**.

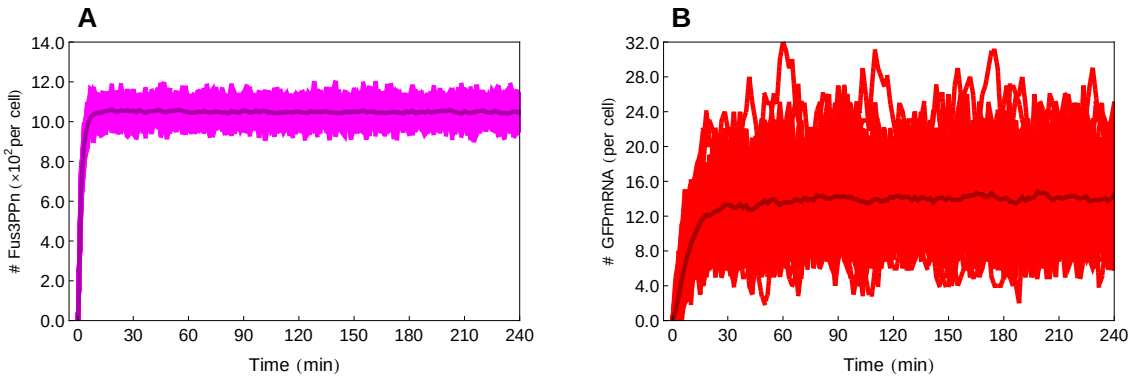


Figure 3.16: Stochastic simulation of the full reporter cell model. We stimulated the reporter cell model with 2.5 nM α-factor. In each case 100 trajectories are plotted over four hours showing the number of **(A)** Fus3PPn **(B)** *GFP*mRNA molecules produced (per cell). The bold grey line in each panel depicts the mean number of molecules for respective compound.

As already mentioned, it was possible to apply QSSA at the receptor level, because the stimulation with constant α-factor resulted in rapid reaching of steady state.
Unfortunately, the sender cells that we model accumulate the α-factor over the simulation time (e.g. **Figure 3.4**), so the reporter cell is exposed to an altering α-factor concentration that prevents the receptor from reaching steady-state. Therefore, the QSSA cannot be applied and the full model of the stochastic cell becomes too time-expensive to simulate when combined into models of logic gates. As a solution, we propose a reduced version of the stochastic model.

3.3.2 Reduced model

The most important and also most noise contributing (see **Results, Noise quantification**) part of the reporter cell model is the GFP expression module (**r34**, **r34deg**, **r35**, **r36**) presented in **Figure 3.17**. Therefore we decide to reduce the stochastic model of the reporter cell to a model including only these four reactions. We use the concentration of the Fus3PPn obtained in deterministic simulations as an input for the model. The with COPASI generated Fus3PPn trajectories were approximated with polynomial functions with the Curve Fitting Tool in MATLAB (Mathworks, Inc.) and then used as input in the stochastic model.

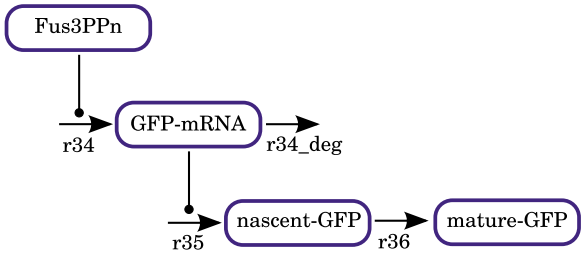


Figure 3.17: Reduced model of the reporter cell: wiring in the system. Each arrow represents a chemical reaction. Lines with dots at the end are modulation reactions.

Chapter 4

Results

Contents

4.1 Deterministic simulations of the gates

As first, we investigate the functioning of the artificially constructed logic gates for different concentrations of input variables employing deterministic simulations. We consider the final amount (at the end of the simulation time) of produced GFP to be the system output for each gate. To this end, we quantify the final concentration of GFP for each model variant. All gates produce biologically feasible concentrations of GFP. We name the cells that produce enough GFP to generate fluorescence as GFP-positive. In order to distinguish between GFP-positive and GFP-negative cells and to evaluate the results we fix a threshold of 4.5 μM of GFP for considering cells as GFP-positive.

All following deterministic simulations were performed with the tool COPASI [27]. The simulations start when the stimulation with corresponding input begins ($t_{start} = 0$) and end after four hours ($t_{end} = 240min$). **Figure 4.1** shows the result of the simulations of the IDENTITY gate model for a range of NaCl concentrations.

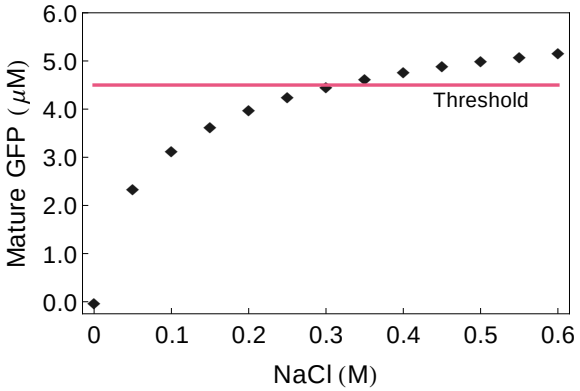


Figure 4.1: Final GFP concentration after four hour stimulation of the IDENTITY gate with different salt concentrations. Increased salt stress results in increased GFP production. The pink line depicts the threshold (4.5 μM) for considering cells as GFP-positive.

The amount of produced GFP by the IDENTITY gate increases when stimulated with increasing salt concentrations. The concentration of produced GFP exceeds the above set threshold when the gate is stimulated with 0.3 M NaCl.
Figure 4.2 presents the result of the deterministic simulations of the NOT gate.

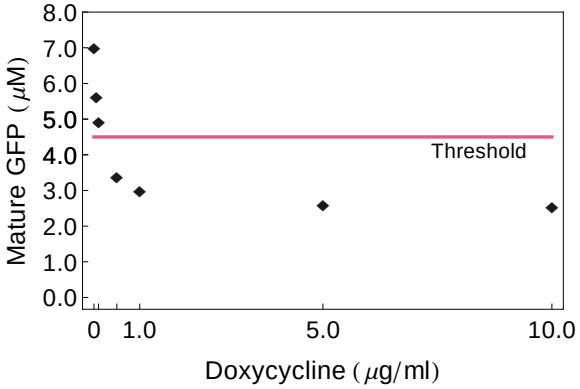


Figure 4.2: Final GFP concentration after four hour stimulation of the NOT gate with different doxycycline concentrations. Increased doxycycline concentration decreases the amount of produced GFP. The pink line depicts the threshold (4.5 μM) for considering cells as GFP-positive.

The NOT gate produces a GFP concentration that exceeds the threshold (i.e. more than 4.5 μM) only for low concentrations of doxycycline (i.e. less than 0.1 $\mu g/ml$). Stimulation with higher doses of doxycycline prevents the reporter cell from producing high levels of fluorescence. Interestingly, the NOT gate allows for basal production of mature GFP as even very high doses of doxycycline still allow the reporter cell to produce ca. 2.5 μM GFP. This phenomenon will be discussed in the last paragraph of this section.
After we have succeeded in constructing gates consisting of one sender and one reporter cell we decided to model more complex gates with two types of sender cells included. **Figure 4.3** presents the result of the deterministic simulations of the 3-cell gates: OR and IMPLIES.

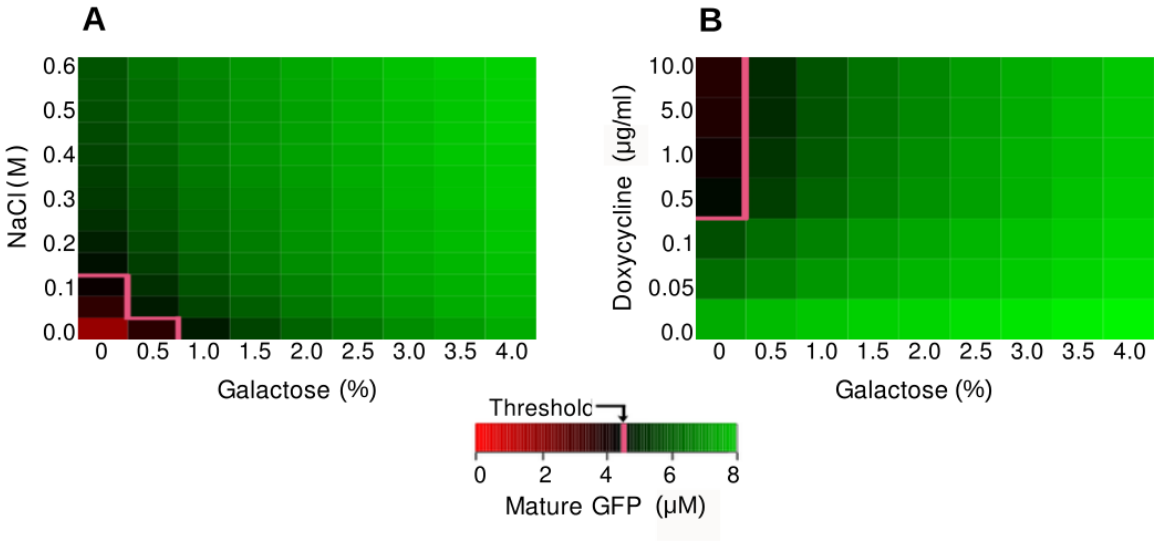


Figure 4.3: **(A)** Final GFP concentration after four hour stimulation of the OR gate with different salt and galactose concentrations. **(B)** Final GFP concentration after four hour stimulation of the IMPLIES gate with different doxycycline and galactose concentrations. The pink line depicts the threshold (4.5 μM) for considering cells as GFP-positive.

The final GFP concentration in 3-cell-gates depends on the degree of stimulation of both sender cells and therefore cannot be presented as a single graph. Thus we use heatmaps to illustrate the behavior of the gates. The amount of produced GFP by the OR gate (**Figure 4.3A**) increases when the system is stimulated with increasing salt or galactose concentrations. Already at low concentrations of these two inputs, the system is able to produce detectable GFP concentration. The result for the IMPLIES gate (**Figure 4.3B**) reveals that the increase in galactose concentration promotes the GFP production. The increase in the concentration of doxycycline in the medium leads to a decrease in the concentration of produced GFP.

As already mentioned in the evaluation of the results of the NOT gate, a significant amount of GFP is being produced even in the presence of high doxycycline concentration. A high doxycycline concentration should in theory fully repress the α-factor production in the dox-cell and ensure that no GFP is being produced by the system. A similar observation can be made for both 3-cell gates. The simulations of the OR gate predict that even if the medium contains neither salt nor galactose, the gate can still produce ca. 1.8 μM GFP. Similarly, the IMPLIES gate can achieve a GFP concentration of ca. 3.7 μM, that is close to the threshold, even if stimulated with high dosage of doxycycline and by simultaneous absence of galactose. These inaccuracies in the functioning of the gates originate from the experimental setup: (i) the gal-cells were originally precultured in galactose and (ii) the dox-cells were precultured without doxycycline in a standard rich medium [1]. The consequence of the preculturing is an internal preaccumulation of the *MFα1*-mRNA and of the PreproAlpha in both gal- and dox-cell before the actual begin of the stimulation. The initial concentrations of these two biochemical species are therefore not equal to zero and it was taken into consideration in our modeling approach. The appropriate initial concentration values that we assigned are listed in the **Appendix**. To be consistent with the experimental procedure the initial concentration of α-factor remains zero

in the model. This is due to the fact that the cells have been always transferred to a fresh medium after the preculturing. As a consequence, the already produced α-factor disappeared. The significance of the preaccumulation of *MFα1*-mRNA and PreproAlpha was investigated further and the results will be presented in the section **Modifications of the system**.

4.2 Sensitivity analysis of the IDENTITY gate

Parameter sensitivity analysis is a useful tool for investigating, how the system variables (e.g. non-constant concentrations of biochemical species) depend on model parameters (e.g. kinetic parameters, initial concentrations of compounds) [36]. To this end, values of parameter sensitivities are calculated. We define sensitivity as the change in the chosen variable upon an infinitesimal change in the particular parameter. It is also useful to scale the sensitivities in order to obtain relative values. These are dimensionless and such sensitivities can be compared with each other. The mathematical formulation for how the variable X_i changes with the change in parameter p_k is given with following equation [37]:

$$S(X_i, p_k) = \frac{\partial X_i}{\partial p_k} \cdot \frac{p_k}{X_i}$$

A standard sensitivity analysis investigates the changes in the steady-state values of variables of interest with respect to the infinitesimal changes in the model parameters. However, not all systems reach steady-state during the course of observation. Moreover, it is sometimes of more importance to know if and how the values of the sensitivities evolve with the time. The final output in our framework (concentration of mature GFP) does not reach the steady-state in all four gates. GFP accumulates in all cases, because there is no degradation included, neither in the experimental framework nor in the model. Considering this fact, we decided to perform time-dependent sensitivity analysis and investigate, how the values of sensitivities change throughout the time evolution of the system.
The time-dependent parameter sensitivity analysis was performed for the IDENTITY gate stimulated with 0.4 M NaCl. The values of sensitivities were calculated for 9 time points (after 15, 30, 60, 90, 120, 150, 180, 210 and 240 minutes of the simulation time) with the tool COPASI [27]. In the following subsection we present the dependence of the output on 33 kinetic parameters, length of the doubling time and density of the culture. We also show, how the accumulation of mature GFP depends on the values of the initial concentrations of the compounds.

4.2.1 Kinetic parameters

Figure 4.4 presents the result of the time-dependent sensitivity analysis for the accumulation of GFP at different time points during stimulation of the IDENTITY gate with 0.4 M NaCl. The heatmap shows that the parameters governing the reactions in the Hog1 pathway of the sender cell (w, K_i, k_{20}-k_{30}) have a very little influence on the GFP concentration during the whole simulation time. The parameters responsible for the production of the α-factor in the salt-cell (k_{31}, k_{32}, k_{33}) have a significant influence on the reporter cell's output, especially during the first hour of model simulation. The module responsible for α-factor production has a positive control over the production of GFP. In other words, if one of the parameters governing the α-factor production increases, the concentration of GFP also increases.

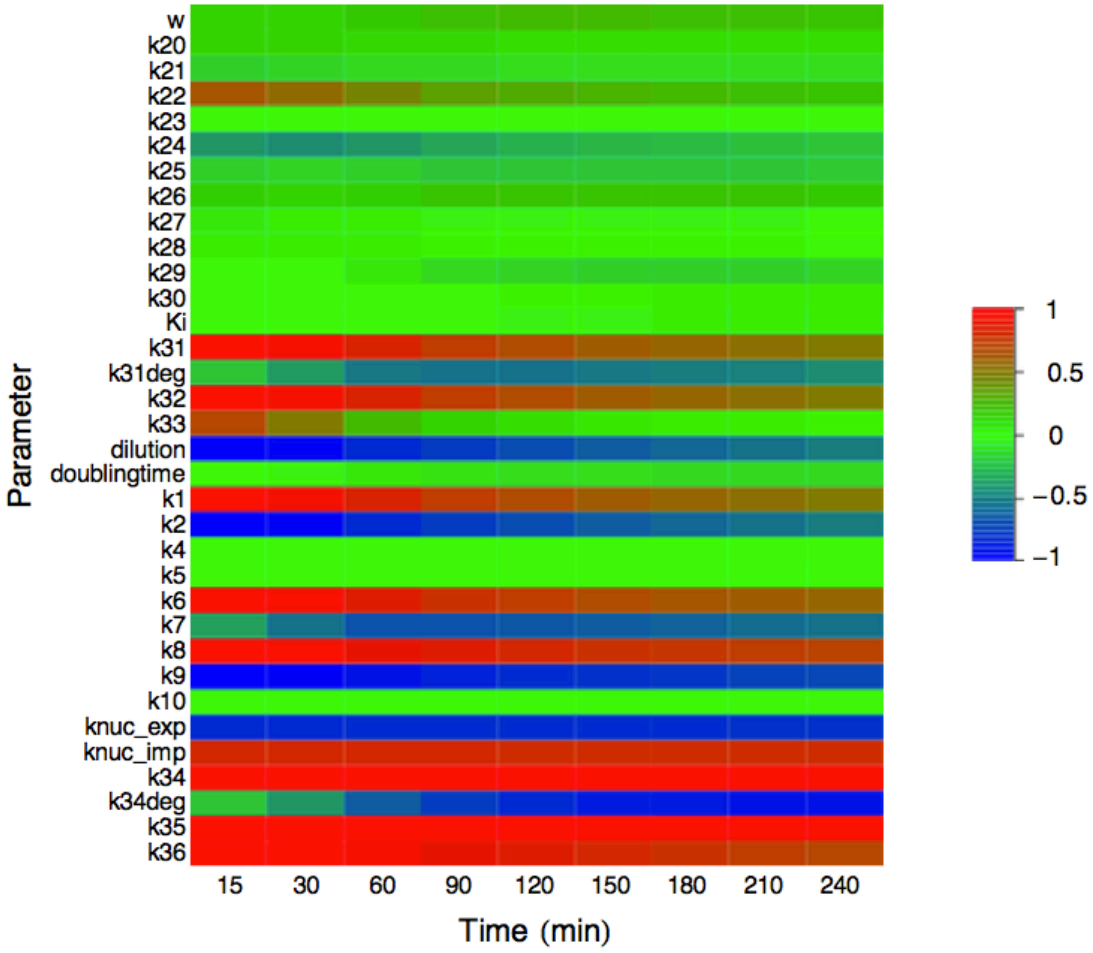


Figure 4.4: Time-dependent sensitivity analysis for GFP accumulated by the IDENTITY GATE stimulated with 0.4 M NaCl. Sensitivities for all kinetic parameters, doubling time and dilution have been calculated for 9 time points during the simulation.

The significance of the length of the doubling time on the GFP production by the reporter cell is negligible at the beginning but increases slightly during the course of the model simulation. The changes in the value of the dilution factor influence negatively the concentration of GFP during the whole simulation time. These two phenomena are discussed in more detail in the section **Modifications of the system** at the end of this chapter.

In the pheromone pathway of the reporter cell, there are three parameters that have a positive influence on the output concentration. These are: the binding of the pheromone to the receptor (k_1), the activation of the Ste5 complex (k_6) and phosphorylation of the cytoplasmic MAPK Fus3 (k_8). As a consequence, the parameters of the reversible reactions: pheromone dissociation from the receptor (k_2), inactivation of the Ste5 complex (k_7) and dephosphorylation of Fus3 (k_9) have a negative influence on the reporter cell's output. The reactions that describe the transport of Fus3 between the cytoplasm and the nucleus have a great impact on the system. The increase in the rate of the export of the Fus3 from the nucleus will decrease the amount of accumulated GFP. The increase in the rate of the import of Fus3 to the nucleus will enhance the GFP production. The parameters of the reactions that are direct precursors of the GFP production (k_{34}, k_{35}) and the rate of GFP maturation itself (k_{36}) have a great positive influence

on the output concentration.

4.2.2 Initial concentrations

The influence of the values of initial concentrations of the compounds (in both salt-cell and reporter cell) on the concentration of mature GFP is presented in **Figure 4.5**. All compounds that occur only during the stimulation are not considered, because their initial concentrations equal to 0.

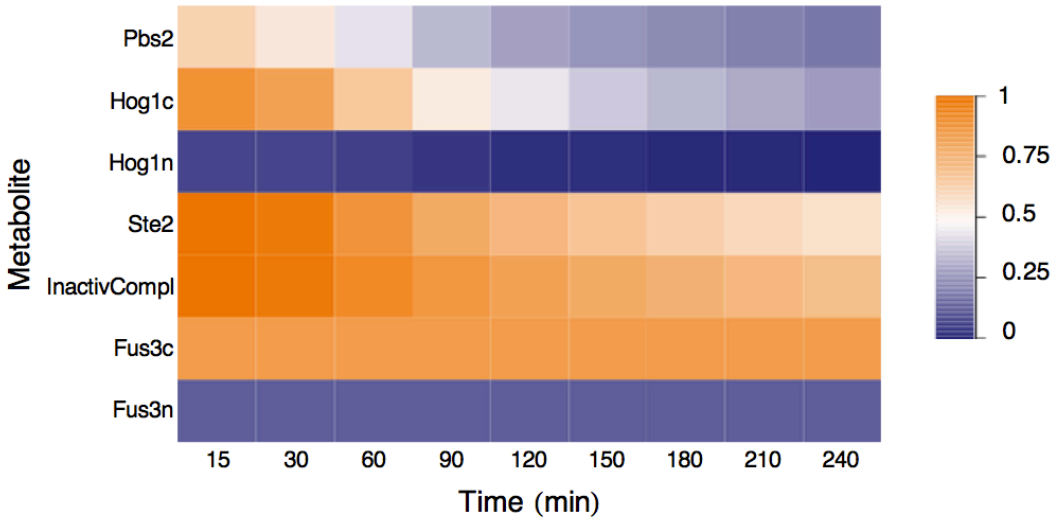


Figure 4.5: Time-dependent sensitivity analysis for GFP accumulated by the IDENTITY GATE stimulated with 0.4 M NaCl. Sensitivities for all non-zero initial concentrations have been calculated for 9 time points during the simulation.

All initial concentration sensitivities values are positive that is a consequence of the cascade-like structure of the whole model and lack of negative regulators. Almost all sensitivities values of initial concentrations decrease with time. Only the sensitivities of initial concentrations of unphosphorylated cytoplasmic and nuclear Fus3 remain the same that is due to the fact, that they reach steady-state almost immediately. The initial concentration of the unphosphorylated cytoplasmic Fus3 has the greatest impact on the system output among all other model variables. It is also interesting that the significance of the initial concentration of nuclear Hog1 is smaller that its cytoplasmic analog. This is likely due to the fact that the MAP kinases can not be phopsphorylated in the nucleus. Therefore, the initial concentration of the nuclear unphosphorylated Hog1 cannot directly contribute to the amount of nuclear phosphorylated Hog1 that drives the process of α-factor production. The same holds for unphosphorylated Fus3 in the reporter cell. Nuclear unphosphorylated form of Fus3 must be first exported from the nucleus to the cytoplasm in order to be phosphorylated. This process is the only way to influence the amount of nuclear phosphorylated Fus3 that directly contributes to the production of GFP.

4.3 Stochastic simulations

As already mentioned in the chapter **Model construction** there was a need for constructing a stochastic model of the reporter cell in order to check how well our models of logic gates reproduces the published population data [1]. In the first of following subsections we quantify the noise in the full (**Figure 3.9**) and reduced (**Figure 3.17**) stochastic model of the reporter cell and justify the reduction of the model. In the second subsection we compare the modeling results of the combined logic gates with the experimental data. All stochastic simulations consider a time scale of four hours.

4.3.1 Noise quantification

Noise in biological systems is a natural consequence of the variation between the cells and between the organisms. It is defined as the variability of the measurements around the mean. It informs us, how distributed are the results of e.g. performance of a signaling pathway. It is crucial for the understanding of the nature of biological system, to know if a particular effect is random or is a result of a precise information transduction by the system. Moreover, it is of high importance to know which components and processes are most noise-contributing to the system and why. In order to find solutions to these questions it is reasonable to quantify the noise. To this end, we calculated the coefficients of variation (CV) using the standard deviation σ and mean μ of the results applying the following equation:

$$CV = \frac{\sigma}{\mu} \cdot 100\%$$

We calculated the coefficients of variation at the nuclear phosphorylated Fus3 (Fus3PPn) level and GFP level in order to be able to distinguish between the noise coming from the pheromone pathway and from the GFP transcription module. The coefficients were calculated at three time points: after 1, 2, and 4 hours in order to follow the time evolution of the noise distribution in the system. As first, we quantified the noise in the full model of the reporter cell stimulated with three different constant concentrations of α-factor (0.5, 2.5 and 5 nM). The mean and the standard deviation were derived from 100 runs of the simulation. The results are presented in **Table 4.1**.

		Coefficient of variation [%] for					
		Fus3PPn			**GFP**		
Model	Variant	1h	2h	4h	1h	2h	4h
Reporter cell	α-factor = 0.5 nM	7.64	7.72	7.65	22.88	16.32	10.33
	α-factor = 2.5 nM	4.25	4.37	3.58	14.71	10.99	7.87
	α-factor = 5.0 nM	3.35	3.73	3.68	15.99	11.62	7.17

Table 4.1: Quantification of the noise in the accumulation of nuclear phosphorylated Fus3 and accumulation of mature GFP in the full model of the reporter cell when stimulated with different α-factor concentrations at three time points (after 1, 2 and 4 hours).

We observe that an increase in the α-factor concentration results in a decrease of noise in the system on both Fus3PPn and GFP levels for most of the time points. At a low concentration of α-factor only a small part of the Ste2 receptor molecules can be occupied that results in

an increase of fluctuation level in the system. The values of CVs for Fus3PPn remain nearly constant during the time. It is due to the fact that Fus3PPn reaches the steady state within the first 15 minutes of the simulation. The values of CVs for GFP decrease with the simulation time. We can think of two explanations for this phenomenon: (i) GFP does not reach steady-state within the time scale investigated and (ii) the continuous increase in the number of GFP molecules reduces the stochastic effects. When the reporter cell is stimulated with low (but constant) concentration of α-factor the main noise in the system originates from the pheromone pathway. However, the contribution to the noise of the GFP production level increases when the cell is stimulated with higher α-factor concentrations.

As next we tried to use the full reporter cell model for the simulations in combination with the sender cells. Our idea was to perform deterministic simulations of the sender cells in order to obtain the time courses of α-factor and then to use them as an input for the full stochastic reporter cell model. Unfortunately, the concentration of α-factor produced by the sender cells alters during the simulation time (**Figure 3.4**, **Figure 3.6**, **Figure 3.8**). As a consequence, it came out that the full model is too time-consuming to simulate when exposed to altering α-factor concentration. Thus, we decided to change our approach and to use the reduced reporter cell model (**Figure 3.17**). In the changed approach, the deterministically obtained time courses for Fus3PPn act as inputs for the reduced model. In other words, not only the sender cell but also the pheromone pathway of the reporter cell is modeled deterministically in the combined models of logic gates.
Before we continue with the noise quantification in the reduced reporter cell model, we present 100 runs of the hybrid model that uses as an input the deterministic time course of Fus3PPn and simulate the next steps stochastically. To show an example, we first simulated the deterministic model of the IDENTITY gate with 0.4 NaCl and obtained the time course for Fus3PPn. Then the time course was approximated with a polynomial function and its derivative was used as the propensity for the production of Fus3PP in the reduced stochastic model. The next steps were simulated in a stochastic manner and **Figure 4.6** presents the results. For simplification, we call the following simulations of the logic gates as 'stochastic', but they are, as explained above, only partially stochastic. This holds also for the next sections and chapters.

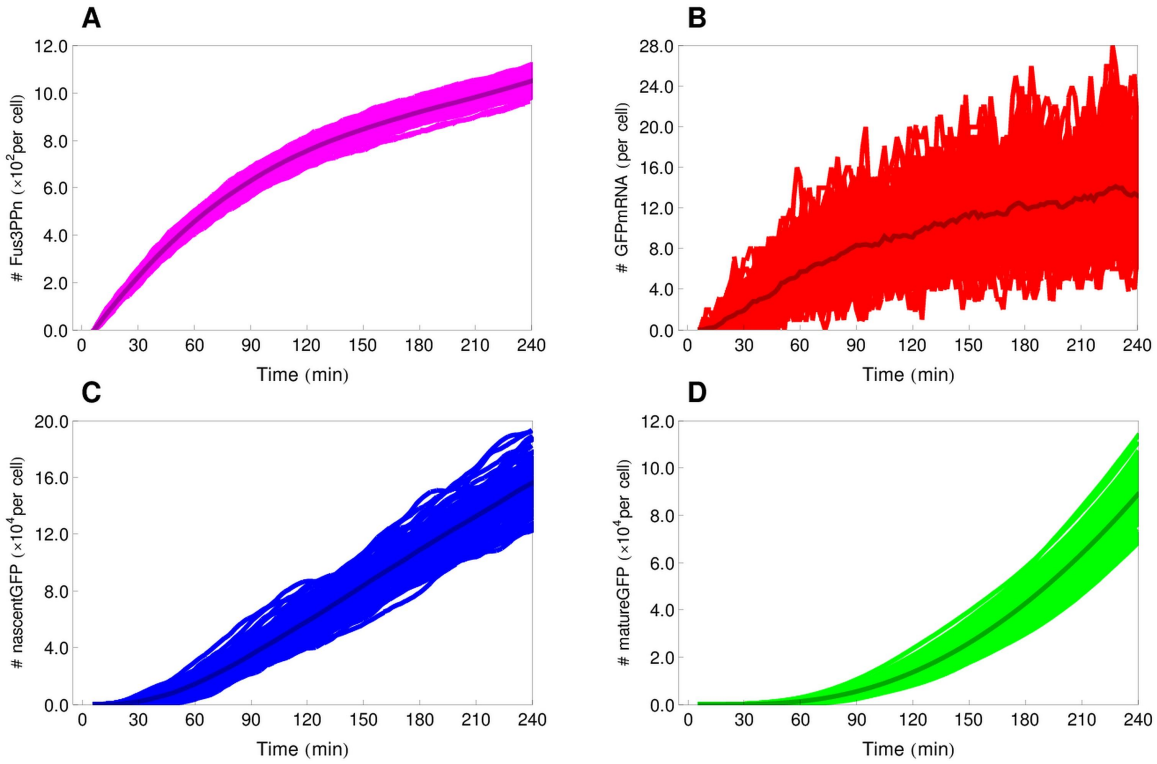


Figure 4.6: Stochastic simulation of the IDENTITY gate stimulated with 0.4 M NaCl with the reduced reporter cell model. For each compound 100 trajectories are plotted showing the number of **(A)** Fus3PPn **(B)** *GFP*mRNA **(C)** nascent GFP and **(D)** mature GFP molecules produced (per cell). The bold grey line in each panel depicts the mean number of molecules for respective compound.

As next, we quantify the noise in the reduced stochastic model. The deterministic concentration of Fus3PPn acts as the input for the stochastic model. Therefore, as already mentioned above, we first obtained the deterministic time courses of Fus3PPn from COPASI . We simulated all four gates for a range of values of corresponding inputs. The coefficients of variation at the Fus3PPn level and GFP level at three time points (after 1h, 2h and 4h) have been calculated. Always 1000 trajectories were used to calculate the mean and the standard deviation. The results are collected in **Table 4.2**.

		Coefficient of variation [%] for					
		Fus3PPn			GFP		
Model	Variant	1h	2h	4h	1h	2h	4h
IDENTITY gate	NaCl = 0.0 M	0.0	0.0	0.0	0.0	0.0	0.0
	NaCl = 0.1 M	6.34	4.62	3.55	47.50	24.00	12.56
	NaCl = 0.2 M	5.36	4.13	3.27	41.16	20.08	11.59
	NaCl = 0.3 M	4.90	3.79	3.14	38.90	19.47	10.64
	NaCl = 0.4 M	4.82	3.79	3.10	37.66	19.06	10.62
	NaCl = 0.5 M	4.52	3.60	3.15	34.68	17.92	10.27
	NaCl = 0.6 M	4.38	3.56	3.09	34.35	18.15	9.93
NOT gate	DOX = 0.0 $\mu g/ml$	3.63	3.20	2.86	26.39	14.59	8.51
	DOX = 0.05 $\mu g/ml$	4.29	3.52	3.05	28.91	16.51	9.78
	DOX = 0.1 $\mu g/ml$	4.58	3.84	3.19	28.86	16.85	10.27
	DOX = 0.5 $\mu g/ml$	5.49	4.79	3.94	32.89	20.72	12.89
	DOX = 1.0 $\mu g/ml$	5.93	5.23	4.24	34.34	21.65	13.30
	DOX = 5.0 $\mu g/ml$	5.81	5.37	4.85	34.05	22.45	14.35
	DOX = 10.0 $\mu g/ml$	6.05	5.61	4.89	35.20	22.52	14.83
OR gate	NaCl = 0.0M Gal=0%	7.16	6.69	5.92	41.45	26.83	17.78
	NaCl = 0.4M Gal=0%	4.15	3.41	2.99	30.01	16.37	9.58
	NaCl = 0.0M Gal=2%	4.20	3.40	2.99	30.12	16.09	9.54
	NaCl = 0.4M Gal=2%	3.68	3.10	2.83	25.15	14.16	8.56
IMPLIES gate	DOX = 0.0 $\mu g/ml$ Gal=0%	3.54	3.18	2.87	23.20	14.55	9.21
	DOX = 0.0 $\mu g/ml$ Gal=2%	3.29	2.98	2.77	21.64	13.37	8.76
	DOX = 10.0 $\mu g/ml$ Gal=0%	4.74	4.56	3.94	27.60	18.51	12.45
	DOX = 10.0 $\mu g/ml$ Gal=2%	3.88	3.36	2.94	23.87	14.05	8.91

Table 4.2: Quantification of the noise in the accumulation of nuclear phosphorylated Fus3 and accumulation of mature GFP in the reduced model of the reporter cell in combination with sender cells. The CVs are calculated for all four logic gates stimulated with corresponding inputs at three time points (after 1, 2 and 4 hours).

For all model variants the coefficients of variation, for both Fus3PPn and GFP level, decrease with the time. It is clearly due to the accumulation of these compounds. In the IDENTITY gate, the increase of the salt concentration promotes the production of α-factor that in turn enhances the activation of Fus3 and, subsequently, GFP production. If many GFP molecules are produced, stochastic effects become less important and the noise level decreases. In the NOT gate, the increase in doxycycline concentration inhibits the α-factor production and thus inhibits the Fus3 activation and GFP production that leads to higher noise levels.

If the OR gate is stimulated with neither salt or galactose, the noise reaches high values for both investigated levels. This is due to a low α-factor and, subsequently, low GFP production. In the IMPLIES gate we observe high noise at both levels only in the absence of galactose and during simultaneous stimulation with doxycycline. In all gate variants, the GFP module contributes much more to the noise than the Fus3PPn production. The noise at the Fus3PPn level ranges from 3 to 5 % that is in agreement with the amount of noise originating from the pheromone pathway in the full stochastic model (**Table 4.1**) and therefore we justify the use the of reduced model for the stochastic simulations which is much less time-consuming. We use the reduced model in order to obtain the population result.

4.3.2 Population result

The available experimental data provide us with information, how many of the reporter cells responded to the stimuli. These were counted as GFP-positive among the cells that were able to respond when stimulated with synthetic α-factor. Our goal was to investigate, if our models of logic gates can reproduce these results. To this end, we simulated 1000 trajectories of mature GFP for each variant of the gate models employing the hybrid models that use the reduced reporter cell model (as input the deterministically obtained time courses for Fus3PPn were used). We counted how many of the trajectories exceed the threshold ($4.5\ \mu M = 78561$molecules) before the end of the four hour simulation. The simulations were performed using the tool CAIN [32]. The results for the IDENTITY and NOT gates are presented in **Figure 4.7**.

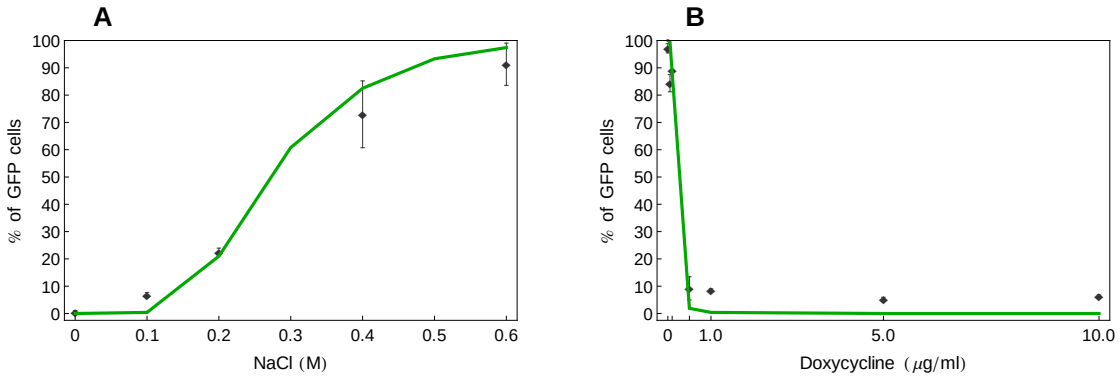


Figure 4.7: Population result for the **(A)** IDENTITY gate and **(B)** NOT gate. The percentage of GFP-positive cells is plotted (green lines) for different input concentrations and compared with the experimental data [1] (black diamonds). As GFP-positive cells we count cells, that exceed the threshold (4.5 μM) before the end of four hour simulation. Statistics is made based on the data set of 1000 stochastic simulations.

The model predictions are in a good agreement with the experimental result for both the IDENTITY and NOT gate. .

We also investigated the functioning of the 3-cell gates models and present the result in **Figure 4.8**.

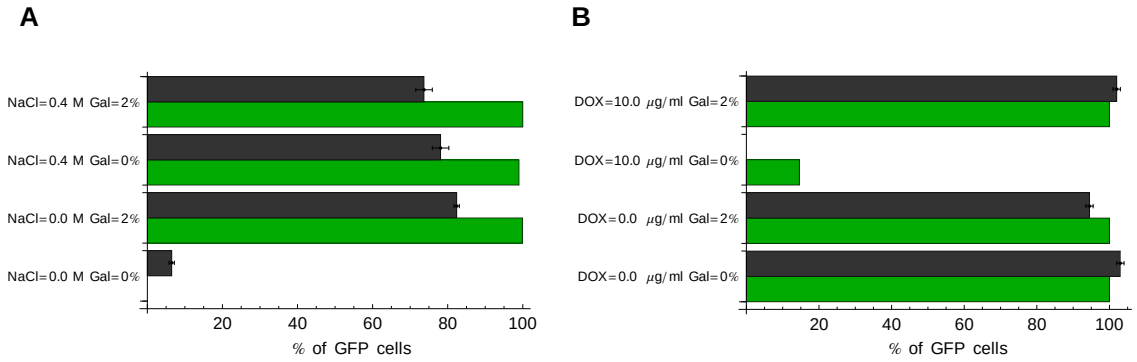


Figure 4.8: Population result for the **(A)** OR gate and **(B)** IMPLIES gate. The percentage of GFP-positive cells is plotted (green bars) for four different input combinations and compared with the experimental data [1] (black bars). As GFP-positive cells we count cells, that exceed the threshold (4.5 μM) before the end of four hour simulation. Statistics is made based on the data set of 1000 stochastic simulations.

The models of OR and IMPLIES gates predict the behavior of the system slightly more digital than the experimental setup. The main divergence is the fact that the model of the IMPLIES gate predicts ca. 15% of GFP-positive cells for the stimulation with doxycycline and in the absence of galactose while the experiment reports no GFP-positive cells. This is due to the preacculumation of *MFα1*-mRNA and PreproAlpha that leads to low production of GFP even if the cells are not stimulated. Despite this minor discrepancies, we assume our models to be good predictors of the experimental behavior of the logic gates.

4.4 Modifications of the system

4.4.1 Gates without preaccumulation of *MFα1*-mRNA and PreproAlpha

As already revealed in the section **Deterministic simulations** the gal-cell can produce small amounts of α-factor even in the absence of galactose as well as the dox-cell can still produce α-factor when stimulated with high doxycycline concentration. These effects are the consequence of the preaccumulation of *MFα1*-mRNA and of the PreproAlpha due to the preculturing. At that point we assumed in our models that the sender and reporter cells are mixed immediately after the end of the preculturing and that the preaccumulated *MFα1*-mRNA and PreproAlpha molecules could not degrade. However, it cannot be ruled out, that there is a delay between the preculturing and start of the stimulation. The half-lives of *MFα1*-mRNA and PreproAlpha are not long (3 min 40 sec and 5 minutes respectively). It is enough to consider a 15-30 minutes long delay due to e.g. the duration of the transfer to a fresh medium, to be able to assume that the compounds can degrade during this time-window. If this is the case, we do not need to consider the preaccumulation of *MFα1*-mRNA and of the PreproAlpha any longer and the initial concentrations can be set to 0. We decided to test this hypothesis and perform the deterministic simulations again.

Figure 4.9 presents the GFP production in the NOT-gate with dox-cells without preaccumulated *MFα1*-mRNA and PreproAlpha.

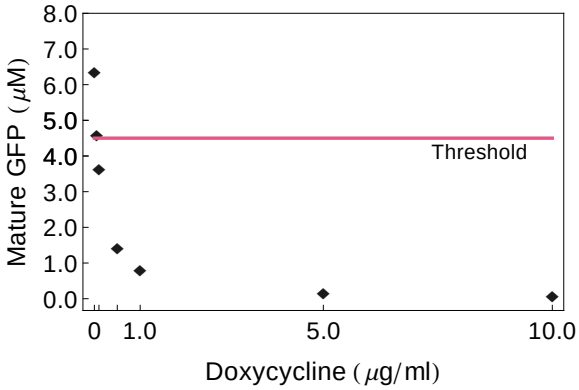


Figure 4.9: Final GFP concentration after four hour stimulation of the NOT gate for different doxycycline concentrations without preaccumulation of *MFα1*-mRNA and PreproAlpha. The initial concentrations of MFα1-mRNA and PreproAlpha in the sender cell have been set to 0. The pink line depicts the threshold (4.5 μM) for considering cells as GFP-positive.

In the NOT gate, the amount of GFP exceeds the threshold only in absolute absence or under a very low concentration of doxycycline (less than 0.05 $\mu g/ml$). The increase in the doxycycline concentration leads to a very fast decrease of GFP levels. There is no GFP being produced for high doxycycline concentrations (more than 1.0 $\mu g/ml$).

We also investigated the consequence of the assumption, that *MFα1*-mRNA and PreproAlpha have both initial concentrations equal to 0 in the OR and IMPLIES gate. In **Figure 4.10** we illustrate the GFP concentration accumulated in the reporter cell when stimulated with different concentrations of salt and galactose (OR gate) or galactose and doxycycline (IMPLIES gate).

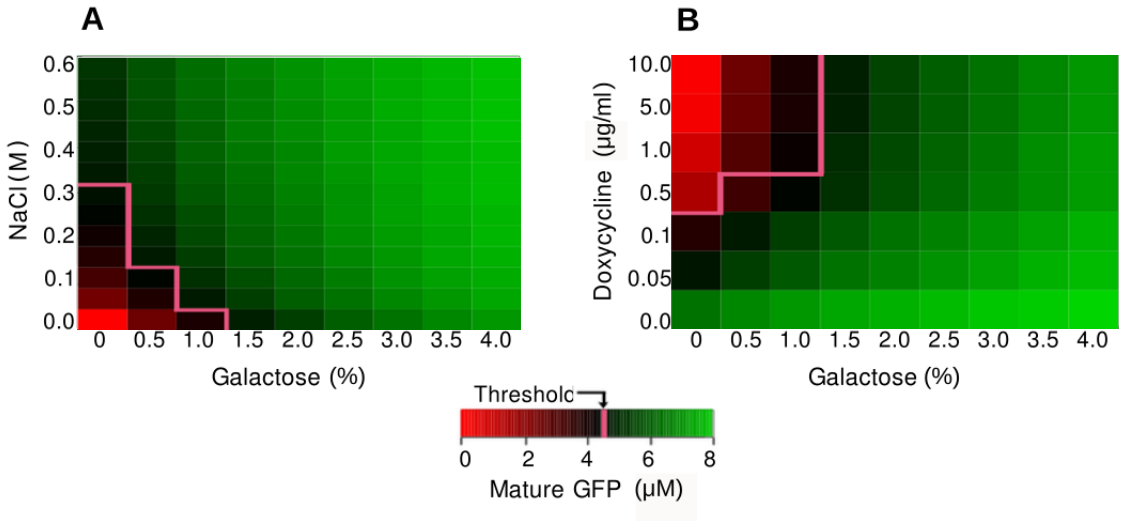


Figure 4.10: Final GFP concentration after four hour stimulation of the **(A)** OR gate for different salt and galactose concentrations and **(B)** IMPLIES gate for different galactose and doxycycline concentrations without preaccumulation of *MFα1*-mRNA and PreproAlpha. The initial concentrations of *MFα1*-mRNA1 and PreproAlpha have been set to 0 in both sender cells. The pink line depicts the threshold (4.5 μM) for considering cells as GFP-positive.

The functioning of the OR gate visibly improved in comparison to previous result (**Figure 4.3**): in the absence of both salt and galactose, no GFP will be produced (**Figure 4.10A**). The ranges of conditions to reach the threshold have also moved to higher input concentrations. Similarly, the behavior of the IMPLIES gate has changed. When the gate is stimulated with doxycycline and in absence of galactose no GFP will be produced (**Figure 4.10B**). The ranges of the conditions for exceeding the threshold are also shifted.
Summing up, the time lap between the preculturing and the start of the stimulation influences the functioning of the gates. The preaccumulation of *MFα1*-mRNA and PreproAlpha prevent the cells from complete termination of the production of the output when not stimulated with galactose or salt (OR gate) or when stimulated with doxycycline and not stimulated with galactose (IMPLIES gate) that anticipates proper performing of the gates.

4.4.2 Density of the culture and doubling time

As we have seen already in the subsection **Sensitivity analysis of the IDENTITY gate** the density of the culture is an important factor for the GFP production. It was also revealed that the doubling time of the culture is a less important parameter. In this subsection we present how the outcome of the system will change if we vary the mentioned parameters using the example of IDENTITY gate stimulated with 0.4 M NaCl. As first we investigate the significance of the density of the culture that was originally set to $5 \cdot 10^6 \frac{cells}{ml}$. In **Figure 4.11** we present deterministic simulations for cultures with two times higher density ($10^7 \frac{cells}{ml}$) and 5 times lower density ($10^6 \frac{cells}{ml}$).

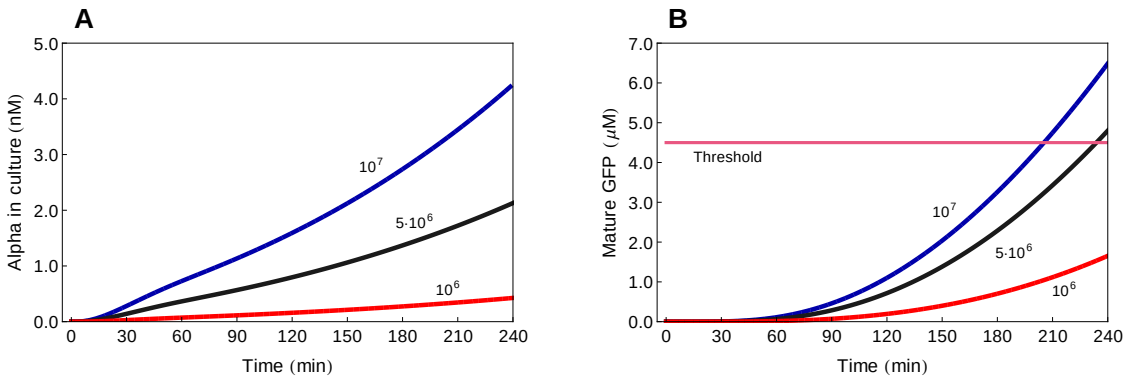


Figure 4.11: Impact of the cell culture density on the concentration of **(A)** α-factor and **(B)** mature GFP. The deterministic simulations illustrate the changes in the concentrations of produced α-factor and GFP for the IDENTITY gate stimulated with 0.4M NaCl for the primary culture density (black lines), for two times higher (blue lines) and five times lower (red lines) density. The pink line depicts the threshold (4.5 μM) for considering cells as GFP-positive.

The result of the simulations at the level of α-factor concentration is obvious. A population consisting of two times more sender cells will produce two times more α-factor because the amount of the produced α-factor is directly proportional to the number of producing cells. Less intuitive is the result at the level of GFP concentration: the population with higher density will

produce less than two times more GFP, but the difference is still significant. The population with lower density produces visibly less GFP (but more than 5 times less) and does not reach the threshold. The reason for the partial loss of information about the α-factor concentration is the switch-like mechanism governing the pheromone pathway (**Figure 3.10, Figure 3.11**).

As second, we investigated the effect of changing the duration of the doubling time of the cell culture. We originally used the doubling time of four hours. In other words, in course of our four hours lasting simulations the number of cells in the culture manages to double. In **Figure 4.12** we present deterministic simulations for cultures with a shorter (2h) and longer (6h) doubling time.

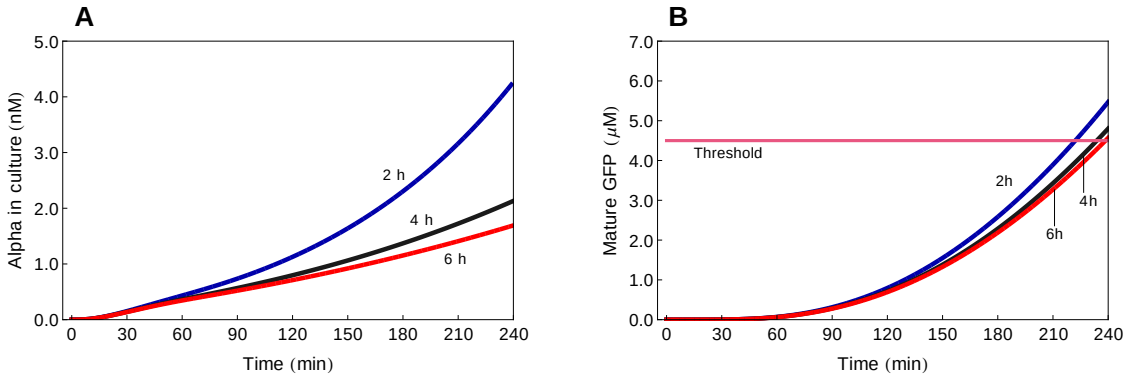


Figure 4.12: Impact of the length of doubling time on the concentration of **(A)** α-factor and **(B)** mature GFP. The deterministic simulations illustrate the changes in the concentrations of produced α-factor and GFP by the IDENTITY gate when stimulated with 0.4 M NaCl for the original doubling time (black lines), for 2 hours shorter doubling time (blue lines) and 2 hours longer doubling time (red lines). The pink line depicts the threshold (4.5 μM) for considering cells as GFP-positive

For the case that the cells double 2 times faster, there is a significant difference in the α-factor concentration in the later part of the simulation. During the first hour the concentration remains almost the same. A similar effect can be observed for the culture with longer doubling time: during the first two hours of simulation there is no difference in the produced amount of α-factor. The differences between the cultures with different doubling times became even smaller at the level of GFP concentration: the culture with shorter doubling time produces not more than ca. 10% more of GFP while the difference for the culture with longer doubling time is almost negligible. Therefore, the threshold can also be reached by a culture with longer doubling time. It should be mentioned, that the effect of altered doubling time will become more significant for simulations that last more than four hours.

Summing up, the analysis of the model revealed that the changes in the density culture have a significant influence on the system performance. It is relatively easy experimentally to alter the density of the culture. Thus, we predict from the model behavior that the levels of produced α-factor and GFP can be easily tuned in an experiment if the culture has an appropriate density.

4.4.3 Less sensitive Ste2 receptor

Until now we have considered the functioning of the gates in terms of boolean logic. Binary systems are most spread and are used most frequently. However, we could here propose an alternative view: a possible design of a device following a three-value logic. The three-value logic operates with three values: 0 (false), $\frac{1}{2}$ (undefined, unknown etc.) and 1 (true)[i]. As presented in the previous sections, salt cell and reporter cell can build an IDENTITY gate. We can extend this gate by adding a second type of reporter cell, that is less sensitive to the α -factor, e.g. reporter cell with a mutated Ste2 receptor.

An Alanine substitution of the 262th residue of the Ste2 receptor increases the value of the dissociation constant (K_d) ca. 2 times so the cell becomes two times less sensitive to the α-factor concentration [38] and will be responding slower to the emerging α-factor gradient. We modeled this mutation by decreasing the parameter value for pheromone binding to the Ste2 receptor: from $k_1 = 8 \cdot 10^{11} \frac{ml}{mmol*s}$ to $k_1 = 4 \cdot 10^{11} \frac{ml}{mmol*s}$. The structure of both models of reporter cells is the same and can be found in **Figure 3.9**. The structure of the extended gate is shown in the **Figure 4.13**.

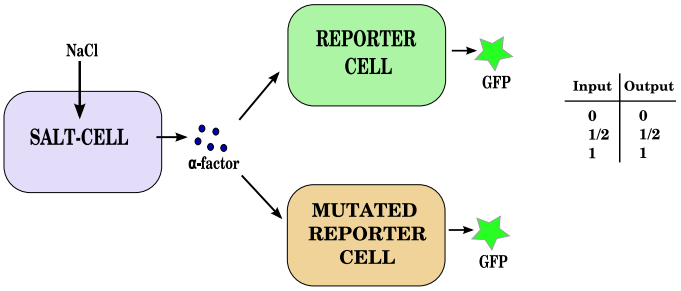


Figure 4.13: EXTENDED IDENTITY Gate: Structure and truth table

After adding the third kind of cell to our model we can test the behavior of the system upon following salt concentration: 0 M, 0.1 M and 0.4 M. We have chosen a simulation time of five hours. **Figure 4.14** presents the deterministic simulations of the GFP accumulation in the original and mutated reporter cell.

When stimulated with no salt, no GFP will be produced for both kind of reporter cells. If stimulated with 0.1 M NaCl, only the original reporter cells produce enough GFP to reach the threshold. However, for higher salt concentration (0.4 M), both populations of reporter cells reach the threshold and produce fluorescence.

[i] The choice is arbitrary. The numerical representation can differ e.g. -1,0,1 or 0,1,2 sets can be employed.

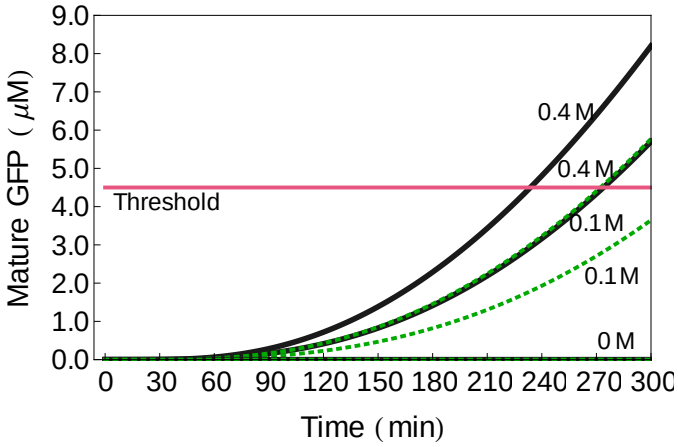


Figure 4.14: Deterministic simulations of the EXTENDED IDENTITY gate. The GFP concentration produced by the wildtype reporter cell (black lines) and by mutated reporter cell (green lines) for three different NaCl concentrations (0.0, 0.1 and 0.4 M) is plotted. The pink line depicts the threshold (4.5 μM) for considering cells as GFP-positive

To strengthen this point, stochastic simulations have been performed and the results are presented in **Figure 4.15**.

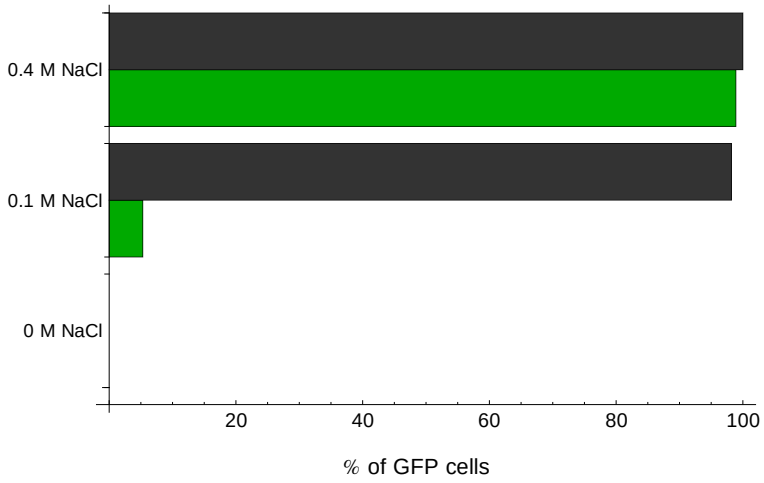


Figure 4.15: Percentage of GFP-positive cells in the EXTENDED IDENTITY gate. The percentage of GFP-positive cells from the wildtype reporter cell population (black bars) and from the mutated reporter cell population (green bars) for three different NaCl concentrations (0.0, 0.1 and 0.4 M) is plotted. 1000 stochastic simulations were run and GFP trajectories were collected. Each trajectory that exceeded the threshold (4.5 μM) before the end of the five hour simulation was considered to represent a GFP-positive cell.

As expected, there are no (**0**) GFP-positive cells when stimulated with no salt (**0**). For a low salt concentration ($\frac{1}{2}$), almost all cells from the population with wildtype receptor can be counted as GFP-positive, while only a few from the population with mutated receptor can produce enough GFP to exceed the threshold ($\frac{1}{2}$). For a high salt concentration (**1**), both reporter cells population types are GFP-positive (**1**). This founding proves that the modified system can act as a three-valued IDENTITY gate.

Chapter 5

Discussion

Contents

Mathematical modeling provides an insight into our understanding of the functionality of biological systems. We employed mathematical modeling to learn more about the performance of synthetically engineered cells and the logic gates constructed of them. As first, we designed and implemented deterministic models of the sender and reporter cells using the published data to obtain the model parameters. We then combined the models of individual cells into logic gates and investigated their performance by running deterministic simulations. Next, we analyzed which of the model parameters have the major influence on the output. These were further investigated. In order to validate the model, we created a stochastic model of the reporter cell and used it in combinations with deterministic models of the sender cells. These hybrid models were used to simulate the population results. Despite minor discrepancies between the model predictions and population data we regard the models as well-optimized and correctly functioning. Finally, we proposed additional modifications of the models and investigated their impact on the performance of the logic gates.

Although the models of the logic gates satisfactory reproduce the experimental data, we have to bear in mind that they were constructed with many simplifications. Below, we discuss these restrictions in detail.

5.1 Evaluation of the model construction

We reduced the complexity of the constructed gates to minimum. Simple structures make the models easier to handle and faster to simulate. Such models are most appropriate for modeling of systems, where the final output of the system is of most interest. A small number of reactions, and hence a small number of parameters prevents also the overfitting of the models. On the other hand, applied simplifications hinder a detailed analysis of the dynamics of some parts of the models. To give an example, the MAPK cascades of the HOG and pheromone pathways are modeled as single steps. This makes the analysis of the dynamics of the individual steps of the cascades impossible.

Beside of simplified structures of deterministic models we applied a broad range of other reductions. We assumed that there is no difference in the doubling rates of the cultures of different

types (all four types of investigated cells: salt-cell, gal-cell, dox-cell and reporter cell have different growth conditions). We also used the same value of translation rate for the α-factor in all sender cell models and for the GFP in the reporter cell model. For the fitting of the models to experimental data, we assumed in the model that maximal phosphorylation levels of Hog1 and Fus3 correspond to a situation where half of the molecules is phosphorylated. Furthermore, we neglected potential changes in the volume of the compartments and model the the volume of the cytoplasm and of the nucleus as constant. We also assumed, that the biochemical species are uniformly distributed inside the compartments. As next, the values of initial concentrations of the compounds were obtained from one source [39] for coherency. Moreover, in the gal-cell we modeled the influence of the presence of galactose in the medium on the functioning of the *GAL1* promoter as a single reaction, neglecting transport and signal transduction processes. The same holds for the modeling of the dox-cell, where the inhibition of the *Tet-Off* promoter is modeled with only one reaction.

We did not consider any stochastic modeling of the sender cells. We justify this with the fact that the α-factor produced by the sender cells is secreted into the medium of well mixed cultures and only its concentration is important. Therefore, we only considered the stochastic modeling of the reporter cell. The major simplification that we made, while constructing stochastic models is the reduction of the full reporter cell model. We neglected the stochastic effects on the sequential steps of the pheromone pathway and approximated the produced noise with one reaction governing the activation of nuclear Fus3. To this end, we fitted the simulated deterministic curves of active nuclear Fus3 with polynomial functions. The derivatives of the resulting equations were used as the propensities for the stochastic activation of nuclear Fus3 in the reduced stochastic model. This simplification reduced the computation time from days to minutes. On the other hand, it can be the major source of errors in the functioning of the gates mainly due to the approximation of the curves that is weighted with an uncertainty.

In summary, we are able to justify all simplifications that we applied during the construction of the models. Despite all of them, we believe that our models are able to be faithful predictors of the behavior of the gates.

5.2 Evaluation of the results

Deterministic simulations The results of the deterministic simulations correspond to our expectations in most cases. The only discrepancy was in the functioning of the OR gate and IMPLIES gate, when stimulated with marginal values of wiring substances. We identified the preaccumulation of PreproAlpha and *MFα1*-mRNA in the gal- and dox-cell to be the reason for this phenomenon. Thus, we proposed in the section **Modifications of the system** altered models of the gal- and dox-cell cells that assume a 30 minute time delay between the preculturing of these sender cells and the mixing with reporter cells. We claim that a time delay between these two events is realistic experimentally and that the experiments that were performed by Regot and colleagues also could have a time delay.

Population result The published population data were used only for the validation of the models of logic gates. The results of the simulations of the models of the gates are in a good

agreement with the experimental data. However, there are some discrepancies, possibly due to following facts: (i) arbitrary set threshold in the model for considering cells as GFP positive (ii) performing simulations with the reduced stochastic reporter cell model and (iii) preaccumulation of PreproAlpha and *MFα1*-mRNA in the sender cells.

Parameter sensitivity analysis The parameter sensitivity analysis is applied to identify the parameters or compounds that influence most significantly the accumulation of the output of the system. We presented the results of parameter sensitivity analysis for the IDENTITY gate stimulated with 0.4 M and found out, that changing the values of parameters governing the synthesis of α-factor in the salt cell and the synthesis of GFP in the reporter cell will result in significant changes in GFP concentration. Parameters responsible for both α-factor and GFP synthesis have positive control above GFP accumulation. In other words, if there is a need for producing more GFP, the translation rates of α-factor or GFP should be increased. It can be performed experimentally by e.g. using a different promoter. However, this procedure requires time to: (i) test the the functioning of the modified construct, (ii) tune it and (iii) run the required controls. Fortunately, we identified also the density of the cultures to have a great impact on the final GFP concentration. Thus, it is simpler to tune the system output via changes in the cell culture density than via genetic re-engineering of the cells.

Three-value logic In the last subsection of **Results** we show how the IDENTITY gate can be extended to a framework functioning as a three-value device. This modification is also easy to achieve experimentally. Although the three-value logic is not often applied in electric circuits it might have an application in biology because of the existence of third possible logic value.

5.3 Outlook

A fast and precise design of biological circuits exhibiting newly programmed behavior is one of the directions in the development of synthetic biology [40]. To this end, modeling and experimental techniques have to be applied in order to design and characterize biological building blocks. These will become parts of a library of modules that can be used to create synthetic circuits. Mathematical modeling not only allows for a detailed characterization of the biological parts but also allows to test modifications and make predictions that decreases the number of experiments to be performed. In this work we investigated in detail the functioning of the models of single cells and gates as well as proposed some modifications. These models can be further elaborated to address many other biological questions related to the models of gates. For instance, it could be investigated, how a potential GFP bleaching or application of diverse fluorescent proteins influences the results. Furthermore, it could be analyzed if there is a possibility to shorten the time needed for producing system output (from few hours in current setup to e.g. one hour). This is tightly connected with the fact that a fluorescent protein is used as a reporter in the cells and its translation and maturation time contribute to response time. Thus, it could be interesting to search for other methods that could be used for fast reporting.
To sum up, we would like to say that synthetic biology can benefit from the computational methods and that the iterative process of designing models, validating them with experiments and finally updating models aids the development of this field.

A.1 Rate equations in the deterministic models

	Reporter cell	
r1	$k_1 \cdot [Ste2] \cdot [AlphaInCulture]$	binding of pheromone to the receptor
r2	$k_2 \cdot [Ste2Ph]$	deactivation of the receptor
r3	v	synthesis of the receptor
r4	$k_4 \cdot [Ste2]$	degradation of inactive receptor
r5	$k_5 \cdot [Ste2Ph]$	degradation of active receptor
r6	$k_6 \cdot [Ste2Ph] \cdot [InactiveComplex]$	activation of the complex
r7	$k_7 \cdot [ActiveComplex]$	deactivation of the complex
r8	$k_8 \cdot [ActiveComplex] \cdot [Fus3c]$	phosphorylation of cytoplasmic Fus3
r9	$k_9 \cdot [Fus3PPc]$	dephosphorylation of Fus3PPc
r10	$k_{10} \cdot [Fus3PPn]$	dephosphorylation of Fus3PPn
r11	$k_{nucexp} \cdot k_{small} \cdot [Fus3PPn] \cdot V_{nuc}$	export of Fus3PPn
r12	$k_{nucimp} \cdot [Fus3PPc] \cdot V_{cyt}$	import of Fus3PPc
r13	$k_{nucimp} \cdot [Fus3c] \cdot V_{cyt}$	import of Fus3c
r14	$k_{nucexp} \cdot [Fus3n] \cdot V_{nuc}$	export of Fus3n
r34	$k_{34} \cdot [Fus3PPn]$	synthesis of GFPmRNA
r34deg	$k_{34deg} \cdot [GFPmRNA]$	degradation of GFPmRNA
r35	$k_{35} \cdot [GFPmRNA]$	synthesis of nascentGFP
r36	$k_{36} \cdot [nascentGFP]$	synthesis of matureGFP
	Salt-cell	
r20	$k_{20} \cdot [Pbs2] \cdot osmostress$[i]	phosphorylation of Pbs2
r21	$k_{21} \cdot [Pbs2PP]$	dephosphorylation of Pbs2
r22	$k_{22} \cdot [Pbs2PP] \cdot [Hog1c]$	phosphorylation of cytoplasmic Hog1
r23	$k_{23} \cdot [Hog1PPc]$	dephosphorylation of Hog1PPc
r24	$k_{24} \cdot [Hog1PPn]$	dephosphorylation of Hog1PPn
r25	$k_{25} \cdot [Hog1PPn] \cdot V_{nuc}$	export of Hog1PPn
r26	$k_{26} \cdot [Hog1PPc] \cdot V_{cyt}$	import of Hog1PPc
r27	$k_{27} \cdot [Hog1c] \cdot V_{cyt}$	import of Hog1c
r28	$k_{28} \cdot [Hog1n] \cdot V_{nuc}$	export of Hog1n
r29	$k_{29} \cdot [Hog1PPc]$	production of internal osmolytes
r30	$\frac{k_{30} \cdot [IntOsmo]}{1+KI*osmostress}$	loss of internal osmolytes
r31	$k_{31} \cdot [Hog1PPn]$	synthesis of $MF\alpha 1$-mRNA ($STL1$ promoter)
	Gal-cell	
r37	$k_{37} \cdot [Gal]$	synthesis of $MF\alpha 1$-mRNA (GAL promoter)
	Dox-cell	
r38	$\frac{k_{38}}{1+Ki \cdot [DOX]}$	synthesis of $MF\alpha 1$-mRNA ($TetOff$ promoter)
	All sender cells	
r31deg	$k_{31deg} \cdot [MFalpha1 - mRNA]$	degradation of $MF\alpha 1$-mRNA
r32	$k_{32} \cdot [MFalpha1 - mRNA]$	synthesis of PreproAlpha
r33	$k_{33} \cdot [PreproAlpha]$	synthesis of Alpha

[i] osmostress=$\max[w \cdot [NaCl] - [IntOsmo], 0]$

A.2 Parameter values in the deterministic models

Parameter	Value	Unit	Reference
k_1	$8 \cdot 10^{11}$	$\frac{ml}{mmol*s}$	The parameters were obtained via fitting to following experimental data sets: 1. Dose-response-curve for receptor occupancy and Fus3 phosphorylation after 15 minutes of exposure to pheromone [41] and 2. Partitioning of Fus3 between nucleus and cytoplasm after pheromone stimulation [42].
k_2	3250	$\frac{1}{s}$	
v[ii]	$6.96 \cdot 10^{-12}$	$\frac{mmol}{ml*s}$	
k_4	$1.84 \cdot 10^{-5}$	$\frac{1}{s}$	
k_5	$2.1 \cdot 10^{-5}$	$\frac{1}{s}$	
k_6	18000	$\frac{ml}{mmol*s}$	
k_7	0.0042	$\frac{1}{s}$	
k_8	$3.2 \cdot 10^{10}$	$\frac{ml}{mmol*s}$	
k_9	680	$\frac{1}{s}$	
k_{10}	0.28	$\frac{1}{s}$	
k_{nucexp}	85.7	$\frac{1}{s}$	
k_{small}	0.5		
k_{nucimp}	16.8	$\frac{1}{s}$	
V_{cyt}	29	fl	We assume the cytoplasm to occupy 50% and the nucleus to occupy 7% [43] of the volume [44].
V_{nuc}	4.06	fl	
k_{34}	$4 \cdot 10^{-6}$	$\frac{1}{s}$	Fitted to *FUS1*-mRNA data [41].
k_{34deg}	0.0214	$\frac{1}{s}$	Recalculated from the half-life time [45].
k_{35}	2	$\frac{1}{s}$	Recalculated from the mean translation rate [46].
k_{36}	$9.625 \cdot 10^{-5}$	$\frac{1}{s}$	Recalculated from the GFP maturation half-time [47].
w	2.39		Fitted to the data for Hog1 phosphorylation [48] and internal glycerol concentration [49].
k_{20}	758	$\frac{ml}{mmol*s}$	
k_{21}	235	$\frac{1}{s}$	
k_{22}	113543	$\frac{ml}{mmol*s}$	
k_{23}	$8.84 \cdot 10^{-5}$	$\frac{1}{s}$	
k_{24}	0.0148	$\frac{1}{s}$	
k_{25}	34.52	$\frac{1}{s}$	
k_{26}	87.84	$\frac{1}{s}$	
k_{27}	5.76	$\frac{1}{s}$	
k_{28}	45.18	$\frac{1}{s}$	
k_{29}	8168.75	$\frac{1}{s}$	
k_{30}	45.18	$\frac{1}{s}$	Fitted to the Hog1 phosphorylation data for *fps1-Δ1* cells (personal communication with Sergi Regot) .
k_{KI}	8168.75	$\frac{1}{s}$	
k_{31}	$1.5 \cdot 10^{-6}$	$\frac{1}{s}$	Fitted to the *STL1*-mRNA data (Elzbieta Petelenz-Kurdziel, manuscript in preparation).
k_{37}	$1.2 \cdot 10^{-11}$	$\frac{1}{s}$	Fitted to *GAL1*-mRNA data [50].
k_{38}	$2 \cdot 10^{-12}$	$\frac{mmol}{ml*s}$	Fitted to the *Tet-Off* promoter data [51].
k_{Ki}	$1 \cdot 10^{7}$	$\frac{ml}{mmol}$	Fitted to the *CLN1*-mRNA data [51].
k_{31deg}	0.0231	$\frac{1}{s}$	Recalculated from *MFα1* half-time [52].
k_{32}	3	$\frac{1}{s}$	Recalculated from the mean translation rate [46].
k_{33}	0.00315	$\frac{1}{s}$	Recalculated from α-factor processing data [53].

[ii] Parameter value ensures that in the absence of pheromone the receptor stays in steady-state.

A.3 Initial concentrations and abundance of species

Species	Initial concentration[iii] [μM]	Molecules per cell
Pbs2	0.1237	2160
Hog1c	0.3405	5948
Hog1n	0.3405	832
Ste2	0.378[iv]	6600
InactiveComplex	0.0385[v]	672
Fus3c	0.406	7090
Fus3n	0.5684	1390

A.4 Initial concentrations of *MFα1*-mRNA and PreproAlpha in gal-cell and dox-cell

Cell	Species	Initial concentration [nM]
Gal-cell precultured in a medium with galactose	*MFα1*-mRNA	0.571
	PreproAlpha	544
Dox-cell precultured in a medium without doxycycline	*MFα1*-mRNA	0.866
	PreproAlpha	825

A.5 Stochastic parameters of the reporter cell models

Parameter	Value [$\frac{1}{s}$]
k_1	45.82
k_2	3250
k_3	0.12144
k_4	0.0000184
k_5	0.000021
k_6	0.000001031
k_7	0.0042
k_8	1.833
k_9	680
k_{10}	0.28
k_{nucimp}	21.423
k_{nucexp}	8.4
k_{34}	0.00000286
k_{34deg}	0.00214
k_{35}	2
k_{36}	0.00009625

[iii] Only non-zero initial concentrations listed. Concentrations recalculated from [39]

[iv] [54]

[v] Ste7 concentration is limiting for the complex concentration

Bibliography

[1] Regot S, Macia J, Conde N, Furukawa K, Kjellen J et al. (2011) Distributed biological computation with multicellular engineered networks. Nature 469: 207–211. (Cited on pages 3, 4, 11, 21, 25, 29, 33 and 34.)

[2] Endy D (2005) Foundations for engineering biology. Nature 438: 449–453. (Cited on page 3.)

[3] Knight TF (2005) Engineering novel life. Mol. Syst. Biol. 1: 2005.0020. (Cited on page 3.)

[4] Bashor CJ, Horwitz AA, Peisajovich SG, Lim WA (2010) Rewiring cells: synthetic biology as a tool to interrogate the organizational principles of living systems. Annu Rev Biophys 39: 515–537. (Cited on page 3.)

[5] Fritz BR, Timmerman LE, Daringer NM, Leonard JN, Jewett MC (2010) Biology by design: from top to bottom and back. J. Biomed. Biotechnol. 2010: 232016. (Cited on page 3.)

[6] Andrianantoandro E, Basu S, Karig DK, Weiss R (2006) Synthetic biology: new engineering rules for an emerging discipline. Mol. Syst. Biol. 2: 2006.0028. (Cited on page 3.)

[7] Elowitz MB, Leibler S (2000) A synthetic oscillatory network of transcriptional regulators. Nature 403: 335–338. (Cited on page 3.)

[8] Bashor CJ, Helman NC, Yan S, Lim WA (2008) Using engineered scaffold interactions to reshape MAP kinase pathway signaling dynamics. Science 319: 1539–1543. (Cited on page 4.)

[9] Zheng Y, Sriram G (2010) Mathematical modeling: bridging the gap between concept and realization in synthetic biology. J. Biomed. Biotechnol. 2010: 541609. (Cited on page 4.)

[10] Anderson JC, Clarke EJ, Arkin AP, Voigt CA (2006) Environmentally controlled invasion of cancer cells by engineered bacteria. J. Mol. Biol. 355: 619–627. (Cited on page 4.)

[11] Ro DK, Paradise EM, Ouellet M, Fisher KJ, Newman KL et al. (2006) Production of the antimalarial drug precursor artemisinic acid in engineered yeast. Nature 440: 940–943. (Cited on page 4.)

[12] Marchisio MA, Stelling J (2011) Automatic design of digital synthetic gene circuits. PLoS Comput. Biol. 7: e1001083. (Cited on page 4.)

[13] Macia J, Sole R (2011) Synthetic Biocomputation: the possible and the actual pp. 29–36. MIT Press proceedings of the european conference on artificial life 2011, ecal 2011 edn. (Cited on page 4.)

[14] Tamsir A, Tabor JJ, Voigt CA (2011) Robust multicellular computing using genetically encoded NOR gates and chemical 'wires'. Nature 469: 212–215. (Cited on page 4.)

[15] Botstein D, Fink GR (1988) Yeast: an experimental organism for modern biology. Science 240: 1439–1443. (Cited on page 5.)

[16] Klipp E, Liebermeister W (2006) Mathematical modeling of intracellular signaling pathways. BMC Neurosci 7 Suppl 1: S10. (Cited on pages 5 and 7.)

[17] Johnson GL, Lapadat R (2002) Mitogen-activated protein kinase pathways mediated by ERK, JNK, and p38 protein kinases. Science 298: 1911–1912. (Cited on page 5.)

[18] Kofahl B, Klipp E (2004) Modelling the dynamics of the yeast pheromone pathway. Yeast 21: 831–850. (Cited on page 5.)

[19] van Drogen F, Stucke VM, Jorritsma G, Peter M (2001) MAP kinase dynamics in response to pheromones in budding yeast. Nat. Cell Biol. 3: 1051–1059. (Cited on page 5.)

[20] Elion EA (2000) Pheromone response, mating and cell biology. Curr. Opin. Microbiol. 3: 573–581. (Cited on page 5.)

[21] Elion EA, Satterberg B, Kranz JE (1993) FUS3 phosphorylates multiple components of the mating signal transduction cascade: evidence for STE12 and FAR1. Mol. Biol. Cell 4(5): 495–510. (Cited on page 5.)

[22] Hohmann S (2002) Osmotic stress signaling and osmoadaptation in yeasts. Microbiol. Mol. Biol. Rev. 66: 300–372. (Cited on page 5.)

[23] Zi Z, Liebermeister W, Klipp E (2010) A quantitative study of the Hog1 MAPK response to fluctuating osmotic stress in Saccharomyces cerevisiae. PLoS ONE 5: e9522. (Cited on page 5.)

[24] Westfall PJ, Ballon DR, Thorner J (2004) When the stress of your environment makes you go HOG wild. Science 306: 1511–1512. (Cited on page 6.)

[25] Klipp E, Liebermeister W, Wierling C, Kowald A, Lehrach H et al. (2009) Systems Biology. Wiley-Blackwell. (Cited on page 7.)

[26] Machado D, Costa RS, Rocha M, Ferreira EC, Tidor B et al. (2011) Modeling formalisms in Systems Biology. AMB Express 1: 45. (Cited on page 7.)

[27] Hoops S, Sahle S, Gauges R, Lee C, Pahle J et al. (2006) COPASI–a COmplex PAthway SImulator. Bioinformatics 22: 3067–3074. (Cited on pages 8, 11, 23 and 26.)

[28] Fogel D, Fogel L, Atmar J (1992) Meta-evolutionary programming. 25th Asiloma Conference on Signals, Systems and Computers. IEEE Computer Society, Asilomar: 540–545. (Cited on page 8.)

[29] Karlebach G, Shamir R (2008) Modelling and analysis of gene regulatory networks. Nat. Rev. Mol. Cell Biol. 9: 770–780. (Cited on page 8.)

[30] Gillespie D (1977) Exact stochastic simulation of coupled chemical reactions. The Journal of Physical Chemistry 81(25): 2340–2361. (Cited on pages 8 and 9.)

[31] Gillespie DT (2007) Stochastic simulation of chemical kinetics. Annu Rev Phys Chem 58: 35–55. (Cited on page 8.)

[32] Mauch S, Stalzer M (2011) Efficient formulations for exact stochastic simulation of chemical systems. IEEE/ACM Trans Comput Biol Bioinform 8: 27–35. (Cited on pages 8, 21 and 33.)

[33] Tamas MJ, Luyten K, Sutherland FC, Hernandez A, Albertyn J et al. (1999) Fps1p controls the accumulation and release of the compatible solute glycerol in yeast osmoregulation. Mol. Microbiol. 31: 1087–1104. (Cited on page 12.)

[34] Chan RK, Otte CA (1982) Physiological characterization of Saccharomyces cerevisiae mutants supersensitive to G1 arrest by a factor and alpha factor pheromones. Mol. Cell. Biol. 2: 21–29. (Cited on page 17.)

[35] Rao C, Arkin AP (2003) Stochastic chemical kinetics and the quasi-steady-state assumption: Application to the gillespie algorithm. J Chem Physics 118: 4999–5010. (Cited on page 21.)

[36] Ingalls BP, Sauro HM (2003) Sensitivity analysis of stoichiometric networks: an extension of metabolic control analysis to non-steady state trajectories. J. Theor. Biol. 222: 23–36. (Cited on page 26.)

[37] Schwacke JH, Voit EO (2005) Computation and analysis of time-dependent sensitivities in Generalized Mass Action systems. J. Theor. Biol. 236: 21–38. (Cited on page 26.)

[38] Lee BK, Lee YH, Hauser M, Son CD, Khare S et al. (2002) Tyr266 in the sixth transmembrane domain of the yeast alpha-factor receptor plays key roles in receptor activation and ligand specificity. Biochemistry 41: 13681–13689. (Cited on page 38.)

[39] Ghaemmaghami S, Huh WK, Bower K, Howson RW, Belle A et al. (2003) Global analysis of protein expression in yeast. Nature 425: 737–741. (Cited on pages 42 and 47.)

[40] Chandran D, Copeland W, Sleight S, Sauro HM (2008) Mathematical modeling and synthetic biology. Drug Discovery Today: Disease Models 5(4): 299–309. (Cited on page 43.)

[41] Blackwell E, Halatek IM, Kim HJ, Ellicott AT, Obukhov AA et al. (2003) Effect of the pheromone-responsive G(alpha) and phosphatase proteins of Saccharomyces cerevisiae on the subcellular localization of the Fus3 mitogen-activated protein kinase. Mol. Cell. Biol. 23: 1135–1150. (Cited on page 46.)

[42] Yu RC, Pesce CG, Colman-Lerner A, Lok L, Pincus D et al. (2008) Negative feedback that improves information transmission in yeast signalling. Nature 456: 755–761. (Cited on page 46.)

[43] Jorgensen P, Edgington NP, Schneider BL, Rupes I, Tyers M et al. (2007) The size of the nucleus increases as yeast cells grow. Mol. Biol. Cell 18: 3523–3532. (Cited on page 46.)

[44] http://yeastpheromonemodel.org/wiki/cell_volume. (Cited on page 46.)

[45] Hyde M, Block-Alper L, Felix J, Webster P, Meyer DI (2002) Induction of secretory pathway components in yeast is associated with increased stability of their mRNA. J. Cell Biol. 156: 993–1001. (Cited on page 46.)

[46] von der Haar T (2008) A quantitative estimation of the global translational activity in logarithmically growing yeast cells. BMC Syst Biol 2: 87. (Cited on page 46.)

[47] Heim R, Cubitt AB, Tsien RY (1995) Improved green fluorescence. Nature 373: 663–664. (Cited on page 46.)

[48] Macia J, Regot S, Peeters T, Conde N, Sole R et al. (2009) Dynamic signaling in the Hog1 MAPK pathway relies on high basal signal transduction. Sci Signal 2: ra13. (Cited on page 46.)

[49] Klipp E, Nordlander B, Kruger R, Gennemark P, Hohmann S (2005) Integrative model of the response of yeast to osmotic shock. Nat. Biotechnol. 23: 975–982. (Cited on page 46.)

[50] Kundu S, Horn PJ, Peterson CL (2007) SWI/SNF is required for transcriptional memory at the yeast GAL gene cluster. Genes Dev. 21: 997–1004. (Cited on page 46.)

[51] Gari E, Piedrafita L, Aldea M, Herrero E (1997) A set of vectors with a tetracycline-regulatable promoter system for modulated gene expression in Saccharomyces cerevisiae. Yeast 13: 837–848. (Cited on page 46.)

[52] Herrick D, Parker R, Jacobson A (1990) Identification and comparison of stable and unstable mRNAs in Saccharomyces cerevisiae. Mol. Cell. Biol. 10: 2269–2284. (Cited on page 46.)

[53] Caplan S, Green R, Rocco J, Kurjan J (1991) Glycosylation and structure of the yeast MF alpha 1 alpha-factor precursor is important for efficient transport through the secretory pathway. J. Bacteriol. 173: 627–635. (Cited on page 46.)

[54] http://yeastpheromonemodel.org/wiki/ste2_num. (Cited on page 47.)

Acknowledgments

First of all, I would like to thank Prof. Edda Klipp for giving me the possibility to write this thesis, for being my supervisor and for many helpful comments and suggestions. I would like to thank Dr. Marta Hoffman-Sommer for all the months of working together on the modeling of logic gates. Especially for her patience, for all the discussions that we had and for many valuable suggestions and inspirations. Special thanks go to Dr. Sean Mauch, the author of the software CAIN for his helpful hints and for the program updates. Furthermore I would like to thank all the members of the group for support and cooperation over the last three years, especially Katarzyna Tyc for many valuable discussions and proofreading of this thesis. Additionally I want to thank my family and my friends for supporting me.

Eigenständigkeitserklärung

Hiermit versichere ich, dass ich die vorliegende Masterarbeit erstmalig einreiche, selbständig verfasst und keine anderen als die angegebenen Quellen und Hilfsmittel verwendet habe.

Berlin, den 29.03.2012