University of West Bohemia
Faculty of Applied Sciences
Department of Cybernetics

BACHELOR'S THESIS

Identification and modulation of the
pheromone response pathway step response

Pilsen 2016 Anna Sosnová

Místo této strany bude zadání práce.

Prohlášení

Předkládám tímto k posouzení a obhajobě bakalářskou práci zpracovanou na závěr studia na Fakultě aplikovaných věd Západočeské univerzity v Plzni.

Prohlašuji, že jsem bakalářskou práci vypracovala samostatně a výhradně s použitím odborné literatury a pramenů, jejichž úplný seznam je její součástí.

V Plzni dne 13. května 2016

Anna Sosnová

Declaration

I hereby declare that this bachelor's thesis is completely my own work and that I used only the cited sources.

Acknowledgements

I would like to thank my supervisor, MSc. Daniel Georgiev, PhD., for his patience, guidance and support.

My thanks also go to all members of Cell cybernetics lab, in particular to Tereza Puchrová, Kateřina Pěchotová and Hynek Kasl for their help and valuable advice.

Abstract

Ste12 protein serves as the key transcription factor that regulates the expression of yeast mating genes after pheromone induction. It is a crucial component of several feedback loops within the yeast pheromone response pathway. This work studies, in silico and in vivo, the replacement of the wild type transcription factor with a synthetic transcription factor within the pathway. Both synthetic and wild type transcription factors were characterized and compared. Pheromone dose response was measured and its vertical shift caused by different transcription factor concentrations was described.

Keywords: dose response, network-free modeling, yeasts, mating, pheromone response pathway, transcription factor, pheromone

Abstrakt

Ste12 protein slouží jako klíčový transkripční faktor, který reguluje expresi kvasinkových párovacích genů po přídání feromonu. Je důležitou součástí několika zpětnovazebných smyček feromonové signální dráhy. Tato práce se zabývá, in silico a in vivo, náhradou přirozeného transkripčního faktoru syntetickým. Oba transkripční faktory byly popsány a vzájemně porovnány. Statická charakteristika feromonové aktivace byla naměřena a změnou koncentrace transkripčního faktoru byl docílen její vertikální posun.

Klíčová slova: statická charakteristika, network-free modelování, kvasinky, párování, feromonová signální dráha, transkripční faktor, feromon

Contents

1 Introduction

For most people yeast are known as organisms used in beer industry and food industry in general. However their potential is much bigger than that. Most people do not know that yeast are nowadays used for production of synthetic insulin that is then used by millions of diabetics worldwide. Several vaccines and other biopharmaceuticals are also synthesized using yeast. Synthetic biology is the key scientific field behind the development of such astonishing technologies. It combines systems biology and genetics with engineering and mathematical modelling. Yeast genome is one of the best described genetic materials. Processes that take place inside yeast cells have been deeply studied and it is anticipated that synthetic biology will significantly influence many future inventions.

This thesis focuses on the terminal messenger protein of the yeast pheromone response pathway that plays an important role during the mating process between yeast haploid cells. Proteins in the pathway interact, create complexes, influence each other by both positive and negative feedback loops. All these processes lead to signal transmission from surface to the nuclei. In practice, this complex pathway can be used for detection of many substances and subsequent response production.

The studied Ste12 protein it the transcription factor required for mating. This work characterizes the behaviour of the wild type Ste12 protein by studying the impact of concentration changes on the dose response. Similarly this work characterizes a synthetic transcription factor that can substitute the wild type Ste12. Mathematical models used to describe both transcription factors are introduced and the resulting hypotheses are experimentally validated.

2 Context

2.1 Biological background

2.1.1 Yeast mating

Yeast Saccharomyces cerevisiae haploid a and alpha cells undergo mating to produce single diploid cells. Each haploid cell produces unique pheromone and receptor corresponding to its mating type. A-cells produce a-factor and Ste2 receptor while alpha cells produce alpha factor and Ste3 receptor.

When the a-cells Ste2 receptor detects presence of alpha factor it initiates projection known as shmoo towards the source of this factor, the alpha cell. Similarly, alpha cell starts to shmoo in the direction of a-cell after detection of a-factor. Preparation for mating and projection of shmoo only occurs after detection of pheromone of the opposite mating type. The information about the presence of pheromone on the surface of the cell is transmitted throught yeast pheromone response pathway to the nucleus where it initiates changes in expression of about 200 genes causing growth arrest, creation of the shmoo and subsequent fusion of the mating partners.

Yeast pheromone response pathway is a system of proteins which transfer the signal through protein-protein interactions. The receptor (Ste2/ Ste3) is part of the G-coupled protein receptor complex. G-protein consists of 3 different units. The alpha subunit is stimulated by the pheromone bound receptor. This stimulation leads to phosphorylation (forwarding a phosphate group PO_4) of GDP (guanosine diphosphate) to GTP (guanosine triphosphate) and subsequent release of subunits G$\beta\gamma$ from the trimeric G-protein.

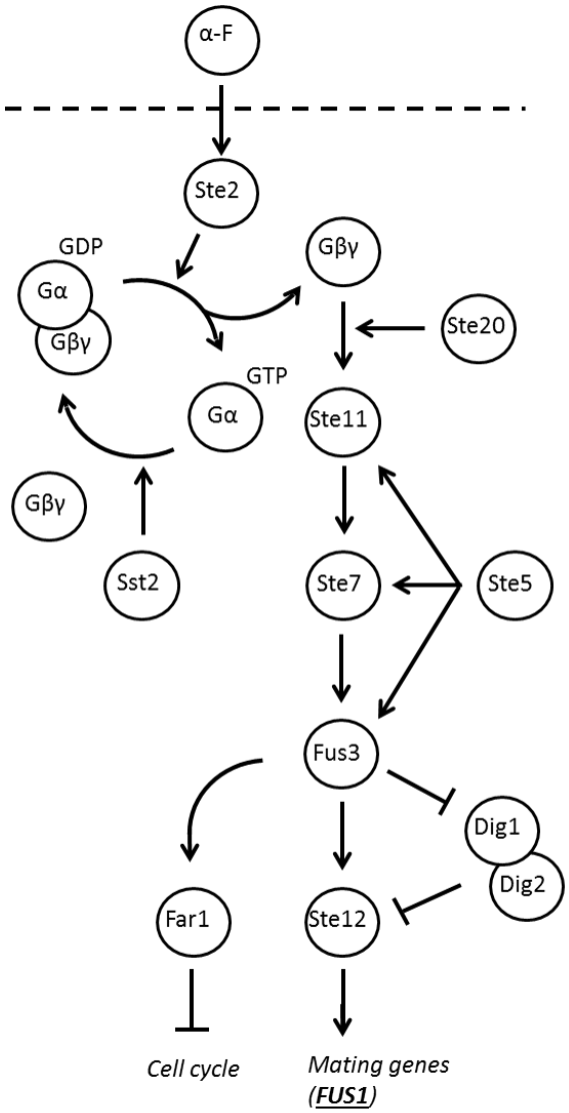


Figure 2.1: Yeast pheromone response pathway [1]. Diagram of fundamental interactions in a-cell. Alpha cell expresses Ste3 receptor instead of Ste2 and is induced by a-factor instead of alpha-factor.

G$\beta\gamma$ binds to two other proteins, Ste20, Ste5 and to Far1-Cdc24 complex. Inactive form of Ste20 present in cytoplasm is activated by its connection to G$\beta\gamma$ and brought near to the plasma membrane where it interacts with other proteins. Ste5, the second target of G$\beta\gamma$ serves as a binding platform and scaffold for several other proteins such as the components of the MAPK cascade.

Mitogen activated protein kinase (MAPK) cascades is common complex in many eukaryotic organisms. It is a set of three sequentially acting protein kinases, MEKK, MEK and MAPK. In the yeast pheromone response path-

way, MEKK is Ste11, MEK is Ste7 and MAPKs are Kss1 and Fus3 [2]. Ste11 is activated by Ste20 via phosphorylation. Ste11 activates Ste7 which afterwards activates Kss1 and Fus3. Target of Fus3 and Kss1 is Ste12/Dig1/Dig2 complex.

There are other proteins that interact with the yeast mating pathway activating, inactivating or regulating its components. However their comprehensive description is beyond the scope of this thesis.

Note that this pathway does not exist in isolation. For instance, several proteins of yeast mating pathway such as Ste12 and proteins of MAPK cascade are also part of the filamentous growth pathway in yeasts that responds to environmental and nutritional stimuli to promote haploid invasive grow and diploid pseudohyphal development.

2.1.2 Ste12

The Ste12 protein serves as the key transcription factor that regulates expression of mating genes after pheromone treatment. A transcription factor is a protein that binds to DNA and regulates gene expression by promoting or suppressing transcription. Transcription is a process during which the DNA template is copied into RNA. It is a first step of protein expression.

In inactive state, Ste12 is part of the Ste12/Dig1/Dig2 complex. Dig2 and Dig1 are negative regulators of Ste12 that keep it inactive when pheromone is not present. After pheromone addition Fus3 and Kss1 phosphorylate Dig1 and Dig2 followed by their unbinding of Ste12 and thus its activation.

Ste12 binds to the genes which contain the pheromone response element (PRE) in their promoters. These are mostly genes responsible for cell mating functions, such as FUS1 and FIG1, which are involved in the process of cell fusion. There are between 1000 and 2000 molecules of Ste12, Dig1 and Dig2 in a single cell and around 200 genes with the PRE element [2].

An important role of active Ste12 is also increasing the copy of certain yeast pheromone response pathway proteins by binding to their gene promoters. These are positive regulators of the pathway STE2, STE4 (beta subunit of G-protein), FAR1 (mediates cell cycle arrest) and FUS3, negative regulators GPA1 (alpha subunit of G-protein), MSG5 (dephosphorylates Fus3), SST2 (stabilize subunits of G-protein) and DIG2. Ste12 is also involved in response to starvation in haploid as it is part of the filamentous grow pathway. During starvation Ste12 is released from Ste12/Dig1/Dig2 complex and forms Ste12/Tec1 complex that afterwards binds to filamentous genes causing flocculation, cell-cell adhesion and cell-substrate adhesion [3].

Ste12 is 688 amino acids long and contain several domains with different function [4]

1. The DNA binding domain (amino acids 1–215) – binds to the PRE element of genes regulated by the Ste12 transcription factor
2. The transcriptional activation domain (amino acids 384-688) – is responsible for activation of both basal transcription induction in absence of pheromone and pheromone induced transcription
3. Induction region (amino acids 216–383) – serves for pheromone induction, amino acids 301–335 are neccesary for pheromone induction

2.1.3 Transcription factor modification

Synthetic transcription factors may be used to redirect the pheromone response pathway signal away from the mating genes to alternate genes of choice.

Pheromone activated synthetic Gal4-Ste12 transcription factor (PAF) combines properties of Gal4 and Ste12 proteins. Gal4 protein mediates expression of galactose inducible promoters such as pGAL1 that have upstream activating sequences (UASGAL) in their sequences to which Gal4 binds. The fact that Ste12 consists of several regions might be used for construction of other synthetic proteins by using domains of Ste12. These proteins then inherit some attributes of Ste12. PAF has previously been constructed through fusion of the DNA-binding domain of Gal4 to the transcriptional activation domain of Ste12 together with most of the induction region of Ste12. This structure enables pheromone induction of galactose inducible promoters characterized by the UASGAL binding sequence. It is important to note that PAF dose not include the DNA binding domain of Ste12 and hence does not activate mating genes containing the PRE binding sequence. The PAF coding sequence is described in [4].

Replacement of the STE12 promoter modulates the production and thereby the concentration of Ste12. The promoter is a regulatory region of DNA located upstream of the protein DNA coding sequence. Extensive collection of both natural and synthetic promoters enables production of various protein concentrations under various conditions. Replacement of the native STE12 promoter with synthetic promoters does not affect the ability of Ste12 to bind to the PRE region of other genes. Nevertheless, with the use of synthetic promoters, Ste12 can no longer affect its own production.

2.2 Problem definition

2.2.1 PAF and Ste12 comparison

Ste12	PAF
Natural	Synthetic
Ste12 DNA-binding domain	Gal4 DNA-binding domain
Self amplification (only with wt promoter)	No self amplification
Positive feedback	No feedback
Activates promoters of mating genes	Activates galactose inducible promoters
Triggers preparation for mating, growth arrest	No growth arrest and no mating preparation

Table 2.1: Differences between Ste12 and PAF behavior

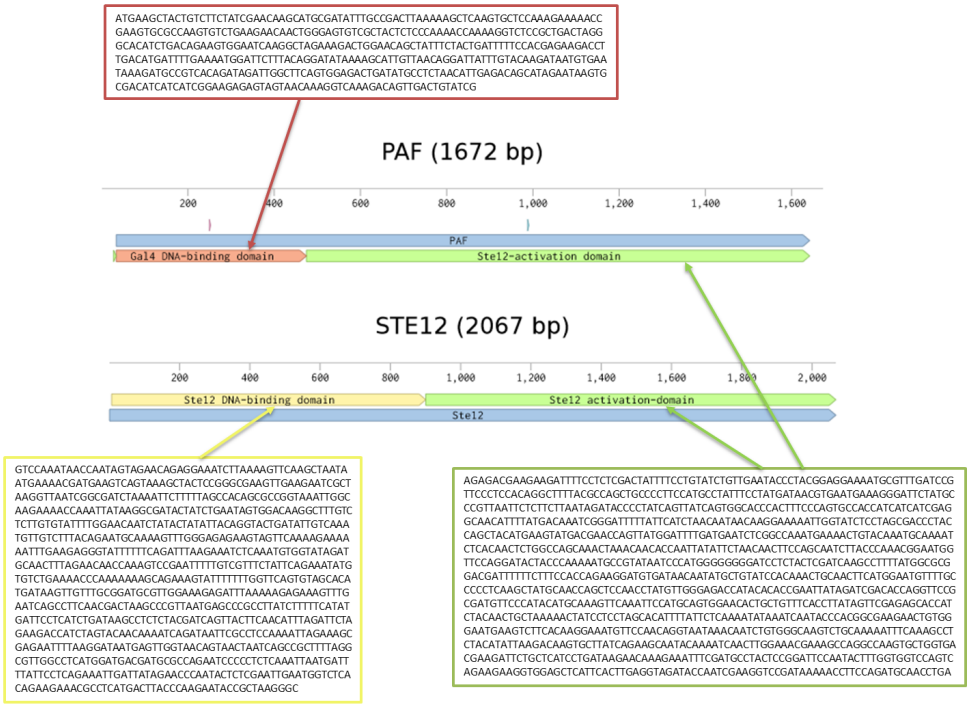


Figure 2.2: Comparison of PAF and STE12 DNA sequence. Both contain Ste12 activation-domain.

Table 1.1 summarizes main differences between Ste12 and PAF. Most significant joint ability of both proteins is their activation in presence of pheromone thanks to their Ste12 activation domain. Both of them are also kept inactive by Dig1/Dig2 proteins in absence of the pheromone.

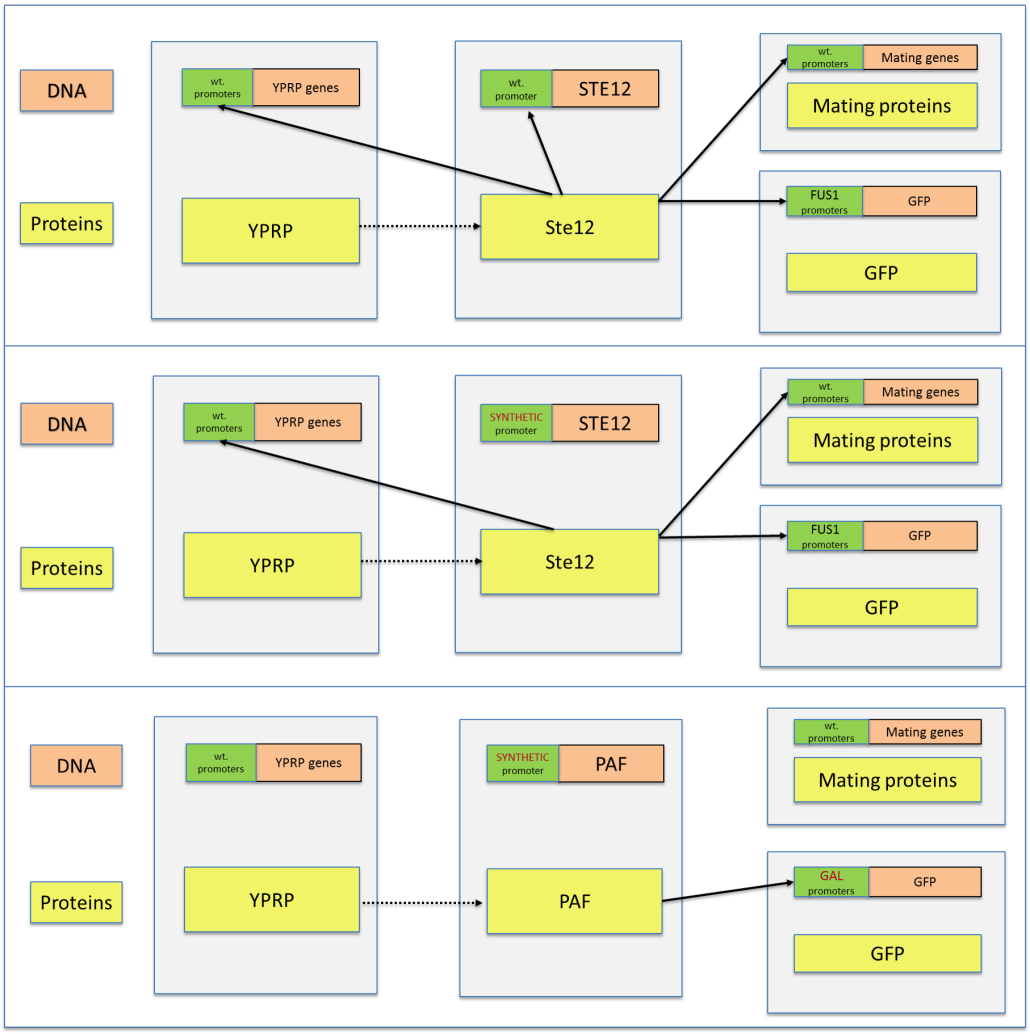


Figure 2.3: Schematic comparison of interactions of Ste12 either regulated by the wild type or synthetic promoter and PAF placed downstream of the synthetic promoter.

Activated Ste12 serves as the transcription factor for mating genes and for genes coding proteins of yeast pheromone response pathway. If placed downstream of the wild type promoter, Ste12 has ability to amplify itself. This self amplification is missing when synthetic promoters are used. However both the positive feedback and activation of mating genes are preserved. PAF does not amplify itself or other components of the pathway. It also

does not trigger the mating response. PAF activates galactose inducible promoters that can regulate almost any chosen gene.

2.2.2 Dose step response

Step response

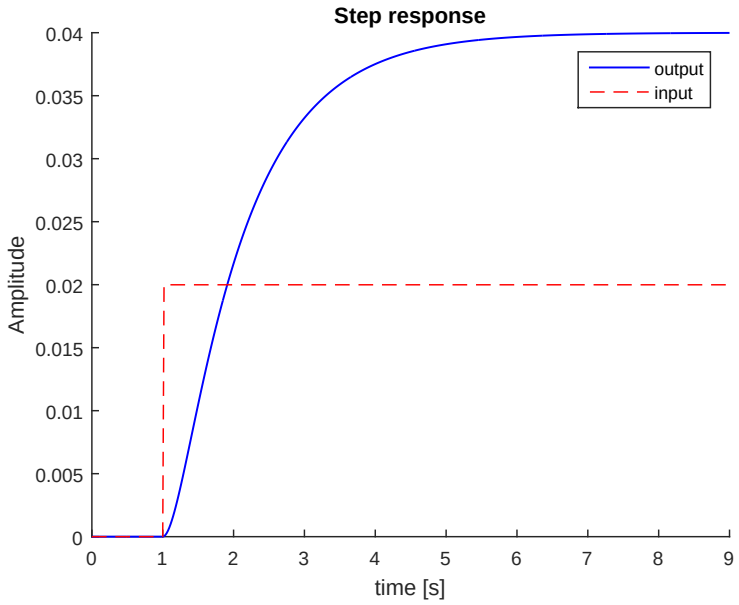


Figure 2.4: Step response

Step response is the time evolution of system output after quick change of input from 0 to a constant value. The step response can be described by quantities related to its time behavior. These are rise time, steady state value and for mechanical systems also overshoot, settling time and ringing [5]. Gene expression step response exhibits the typical characteristics of a second-order system. There is slight delay at the beginning of the production caused by transcription, translation and protein formation. However this delay is often hardly visible since its significantly shorter then the rise time.

Dose response

Dose response curve represents steady stat input-output relationship between dose and response. In this thesis the dose is in the form of a series of different concentrations of alpha factor or alpha cells. The response is stable state of green fluorescence in biological experiments and protein concentration in simulation results. Unlike step response dose response does

not investigate development of the output in time. More likely compares doses after certain time.

The dose response curve is in the form of a **Hill function**.

$$A = V_{max} \frac{A^n}{A^n + K^n} \tag{2.1}$$

V_{max} is maximal value of the response

A is a variable of response

K is constant representing amount of the dose required for 50% activation of the response

n is Hill coefficient which determines the steepness of the curve. Higher values of n indicate quick activation process. The Hill function has form of 0-1 switch for large n.

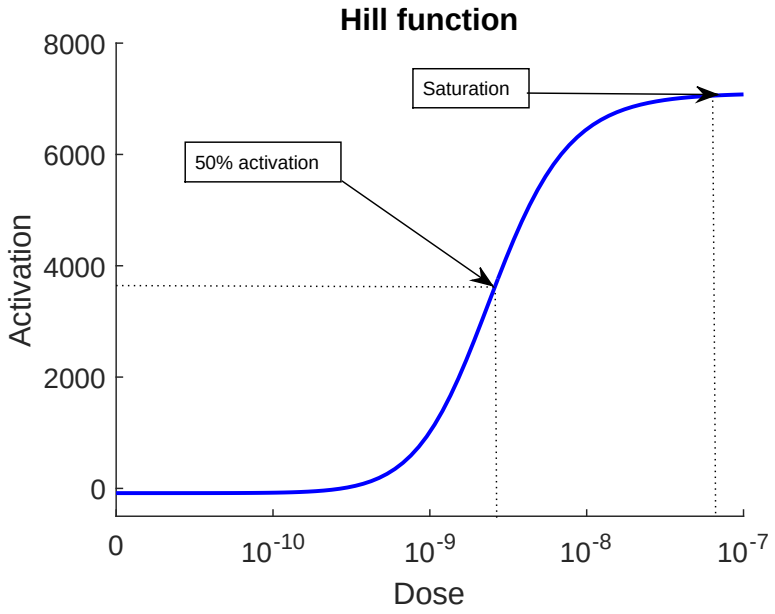


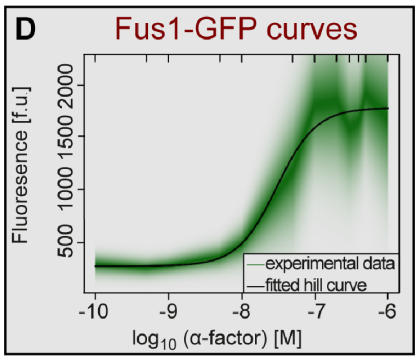
Figure 2.5: Typical Hill function and its key characteristics

Commonly monitored parts of Hill functions are:

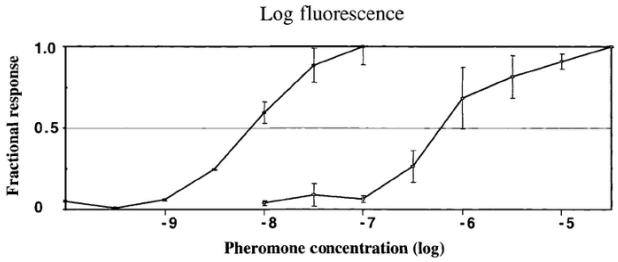
- 50% activation (K) - amount of the dose required for 50% activation (activation threshold)
- Saturation (V_{max}) - maximal achievable value of the response

2.2.3 Dose response variability

For graphic comparison of different constructs behaviour dose response of step responses stable state are widely used in this thesis. Figure 2.6 presents dose responses of pFUS1-GFP reporter plasmid that is directly influenced by Ste12 activation.



(a) Fus1-GFP dose response - in [6], Fus1-GFP dose response in wild type yeast strain BY4741 was experimentaly measured on microscope. Saturation was reached at dose 10^{-6}.



(b) Fus1-GFP dose response - in [7], Fus1-GFP dose response in wild type yeast strain yVT231 (right dose response) was experimentally measured using cytometer. Saturation was not reached.

Figure 2.6: Wild type Ste12 characterization obtained from literature.

Although the constructs used in these measurements were identical, obtained data show significant differences. These are caused by different yeast strains and used methods. Therefore variations among results are expected. Especially during the comparison of in vivo and in silico experiments.

3 Problem solution

3.1 In silico experimental design

Simulations are widely used to predict behavior of biological systems. There are several deterministic and stochastic methods available. Deterministic approach requires solving of differential equations in order to obtain functions describing evolution of all molecule species in time. Frequently all equations include combinations of all existing species and their number is often infinite. Thus, deterministic approach is usually not suitable for simulations of most biological systems due to their complexity.

Stochastic simulations are usually based on numerical solution of the chemical master equation of the system, otherwise referred to as the Kolmogorov forward equation for the Markov jump chain underlying the chemical reaction network.

In order to predict the behavior of the monitored proteins, the Yeast Pheromone Signaling Model developed in [8] was used. This complex stochastic rule-based model was written in BioNetGen language and describes the whole yeast pheromone response pathway.

BioNetGen is a model description language for constructing and simulating rule-based models of biochemical systems [9].

General BNGL file comprise of several sections.

- **Parameters** define particular numerical values that might be used in rule definitions, molecules declaration etc. BioNetGen enables to interactively modify parameters through Matlab.
- **Molecule types** section serves for definition of all molecules, their possible states and binding options.
- **Seed species** defines initial state and number of molecule of the species.
- **Observables** section is a list of species which values are recorded during the simulation. They usually correspond to the data obtained by real experiments.
- **Reaction rules** is a block of rules defining creation and degradation of molecules, their binding and unbinding, state changes.

All possible simulation methods (ODE,SSA,NFsim) use same syntax. This allows to choose the most convenient one. SSA and ODE requires network generation in contrast of NFsim which requires only creation of the XML file.

3.1.1 Yeast Pheromone Signaling Model

Yeast pheromone signaling model represents the yeast pheromone response pathway from G-coupled protein receptor to Ste12 transcription factor. Both negative and positive regulators such as Sst2 are included (see section 2.1.2 Ste12 for their description). Model comprises of approximately 230 rules describing gene expression, degradation, interaction or state transformation. Used rates were either directly observed in yeast (7%), previously used in other models (68%) or estimated (25%) [8]. The model is written in BioNetGen language and network free simulations are recommended due to high complexity.

3.1.2 Mathematical model

Original model was modified to describe the behavior of PAF or Ste12 proteins.

For simulations of **Ste12 transcription factor** reactions describing expression (3.2) and degradation (3.3) of green fluorescent protein (GFP) were defined. GFP gene used in this particular case is under control of FUS1 promoter which included PRE binding region for Ste12 protein (3.1).

$$Ste12 + FUS1p \underset{k_{bs}}{\overset{k_{fs}}{\rightleftharpoons}} FUS1p^* \tag{3.1}$$

$$FUS1p^* \overset{k_{es}}{\rightarrow} FUS1p^* + GFP \tag{3.2}$$

$$GFP \overset{k_d}{\rightarrow} 0 \tag{3.3}$$

Where k_{fs} , k_{bs}, k_{es} and k_d are reaction rates specific for each reaction. $Ste12$ is active Ste12 protein, $FUS1p$ is FUS1 promoter, $FUS1p^*$ is active FUS1 promoter and GFP is green fluorescence protein.

Similarly to the Ste12 case, reactions for **PAF** binding and unbinding to GAL1 promoter were specified. Interaction of PAF and GAL1 promoter (3.4) leads to production of GFP (3.5).

$$PAF + GAL1p \underset{k_{bp}}{\overset{k_{fp}}{\rightleftharpoons}} GAL1p^* \tag{3.4}$$

$$GAL1p^* \xrightarrow{k_{ep}} GAL1p^* + GFP \tag{3.5}$$

$$GFP \xrightarrow{k_d} 0 \tag{3.6}$$

Where k_{fp} , k_{bp}, k_{ep} and k_d are reaction rates specific for each reaction. PAF is active PAF protein, $GAL1p$ is GAL1 promoter, $GAL1p^*$ is active GAL1p promoter and GFP is green fluorescence protein. GFP degradation reaction is the same as in the previous case.

3.1.3 Computer model

This section underlines adaptation of the original model to PAF and Ste12 models implemented in **RuleBender**. RuleBender is a free software for transparent development of rule-based models [10]. It provides an editor for construction, visualization and direct BioNetGen simulation of the model.

The original model does not include rules for RNA transcription and proteins translation which would add delay in proteins production in real system. The duration of these processes is much shorter than other processes performed during simulation. Therefore these rules are left out.

```
Ste12(dna) + Fus1_gene(promoter) -> Ste12(dna!1).Fus1_gene(promoter!1) 2.145e-05
Ste12(dna!1).Fus1_gene(promoter!1) -> Ste12(dna) + Fus1_gene(promoter) 0.03
Ste12(mapk,dig1,dig2,dna!1).Fus1_gene(promoter!1) -> Ste12(mapk,dig1,dig2,dna!1).Fus1_gene(promoter!1) + GFP() 15
GFP() -> Trash() 0.00139
```

Figure 3.1: Deterministic reaction 3.1, 3.2 and 3.3 rewritten to BioNetGen language rules in RuleBender

Both rates for reaction 3.1 were considered same as for all other interactions of Ste12 with promoters used in the model.

Degradation period of the GPF protein is expected to be dependent on the length of the yeast cell cycle (2 hours). Rule for GFP degradation is the same for both Ste12 and PAF models. Equations used for calculation of the degradation rate is [11]

$$k_d = \frac{ln2}{t_{\frac{1}{2}}} = \frac{ln2}{3600} = 0.00193 \tag{3.7}$$

Rate for GFP production was estimated to be the same as the FUS3 production rate wherein the following caveats are in order:

1. Accurate value may differ in specific contexts.
2. Authors of the model also estimated several rates for synthesis reactions in range $10^{-1} - 10^{1} s^{-1}$.
3. The production rate of GFP protein is similar to production rates of other proteins caused by Ste12-promoter binding.
4. Computed GFP degradation rate is equal to estimated Fus3 degradation rate used in original model.
5. There should be thousands of GFP molecules.

The GFP protein undergoes several morphological changes after translation. This maturation process may take couple minutes or even an hour depending on the protein variant. Simulation that includes this process is time consuming. Therefore the concentration of GFP protein exactly after translation is taken to be the system output.

Adaptation of the original model for **PAF simulation** is more extensive. Behavior of PAF is examined in an Ste12 knockout strain (strain lacking the STE12 gene). That's why all rules that included Ste12 were deleted.

Since experimental data describing detailed behaviour of PAF is very sparse, both of the rates for PAF-GAL1 interactions rules were kept same as in the Ste12-FUS1 case. Rate of rule for GFP production regulated by GAL1 promoter was considered same as in FUS1-GFP case .

```
PAF(dna) + Gal_gene(promoter) -> PAF(dna!1).Gal_gene(promoter!1) 2.145e-05
PAF(dna!1).Gal_gene(promoter!1) -> PAF(dna) + Gal_gene(promoter) 0.03
PAF(mapk,dig1,dig2,dna!1).Gal_gene(promoter!1) -> PAF(mapk,dig1,dig2,dna!1).Gal_gene(promoter!1) + GFP() 15
GFP() -> Trash() 0.00193
```

Figure 3.2: Deterministic reaction 3.4, 3.5 and 3.6 rewritten to BioNetGen language rules in RuleBender

Both Ste12 and PAF have ability to interact with Dig1, Dig2, Kss1 and Fus3 proteins. Rules for this interaction containing Ste12 were included in original model. They were all kept exactly the same for PAF.

The main difference between Ste12 and PAF models is inability of PAF to bind to the promoters of yeast pheromone response pathway gene. So the rules for increased proteins production after pheromone treatment are missing.

3.1.4 NFsim algorithm

Stochastic NFsim network-free simulation algorithm for biochemical models was used to simulate behavior of both PAF and STE12 in RuleBender. Network-free simulators do not need to enumerate whole state space of possible species combinations. Only the current state and possible reactions are taken into consideration. This approach highly simplifies the process and usually requires less time and memory. NFsim is integrated into BioNetGen simulator which uses standard deterministic ODE or the stochastic SSA methods. This allow scientists to define same rules using same structure (BioGenNet language, BNGL) for all three methods and then select the proper one.

NFsim algorithm assumes that molecules are randomly distributed throughout the volume and that the whole system is in thermal equilibrium at absolute temperature [12]. Following description summarizes algorithm published in [13].

- Initialization

  At initialization all object representing molecules, rules and functions are created according to the given BNGL file transformed into XML format with structure defined by NFsim authors.

- Step 1 – time sampling

  The waiting time before the next-rule generated is sampled with the same approach as used at SSA algorithm. Propensity (probability that the particular reaction R will occur within [t,t+dt]) of each reaction is calculated and time before the next event is calculated as

  $$\tau = -\frac{1}{r_{tot}} ln(\rho_1) \tag{3.8}$$

  Where r_{tot} is sum of all propensities and ρ_1 is random number from the uniform distribution of interval (0,1).

  Propensity of monomolecular reaction is numerically equal to the reaction rate constant k of conventional deterministic chemical kinetics. For a bimolecular reaction propensity is equal to k/Ω, where Ω is number of molecules of the environment [14].

- Step 2 – next rule selection

The method for selection of the next executed rule is analogous to the SSA direct method. Smallest J satisfying equation

$$\sum_{j=1}^{J} r_j > r_{tot}\rho_2 \tag{3.9}$$

is found. ρ_2 is once again random number from the uniform distribution of interval (0,1). The rule which propensity was the latest added to the sum is then applied.

- Step 3 – reacting molecules selection and validation of their use

  All reactants have an equal chance to be selected. However after their selection it is necessary to verify that they comply with all given constrains (e.g. binding compatibility, state). If not, reaction is not applied and simulation returns to step 1. This step is required due to possible definition of different molecule states.

- Step 4 – reactants transformation and state update

  Selected reactants are transformed according to the selected rule. Reactant list is updated to define their ability to participate in new set of rules.

- Step 5 – update of propensities

  Finally, propensities of all rules are updated based on the current state of the system.

3.2 In vivo experimental design

In following section experiments performed in vivo are presented. Green fluorescence indicating the activation of monitored Ste12 or PAF proteins was measured in 20 minutes intervals until the steady state was reached. Measurements were performed on a plate reader that includes temperature regulation to ensure best conditions for the cells and allows simultaneous measurement of fluorescence and cell optical density (OD).

Two methods of induction were used.

1. **Alpha factor induction** is widely used method based on simple addition of alpha factor to the a-cells sample. Concentrated and purified alpha factor is provided by biological companies. Advantage of this method is knowledge of precise alpha factor concentration. However

this method is quite expensive due to high prize of alpha factor stocks. This allows to perform only few experiments with limited range of alpha factor.

2. **Alpha cells induction** is based on addition of alpha cells to a-cells sample. Alpha cells naturally produce alpha factor which is detected by a-cells and leads to activation of yeast pheromone response pathway. This method is considered more natural since it is based on regular yeast mating initial process. However the amount of alpha factor produced by the cells can be only estimated. In this thesis alpha cells of known OD were added to the sample assuming that the higher number of alpha cells was added the higher concentration of alpha factor was present in the sample. OD measured during the experiment had to be divided in ratio a-cells:alpha-cells. Only the OD of a-cells was used in data normalization. The natural fluorescence of alpha cells is much lower than the fluorescence of a-cells after activation. Therefore the measured fluorescence was attributed only to the a-cells. Low cost of this method allows more repetition of the measurement.

Green fluorescent protein (GFP) exhibits green fluorescence when exposed to the light in the blue to ultraviolet range [15]. It is commonly used in synthetic biology as a reporter of expression. In experiments sequence coding GFP was placed upstream GAL1 and FUS1 promoter. These are induced by activated PAF (GAL1) or Ste12 (FUS1) proteins. Concentration of GFP is therefore directly dependent on number of activated proteins.

The optical density (OD) value represents amount of light absorbed by the sample and is used for estimation of cells concentration in the sample. Measurement of the sample at a wavelength 600 nm (OD600) was performed. Obtained OD data are then used for normalization of the fluorescence.

To compare behaviour of Ste12 and PAF at different copy number their coding sequences were placed downstream of different **4 promoters**. pADH1 is promoter naturally present in yeast cells. It is one of the strongest promoters. pLAC3, pLAC13, pTET20 are synthetic promoters design in [16]. The characterization of promoters strength performed in our laboratory showed that pTET20 is stronger than pLAC13 which is stronger than pLAC3 however the difference is not so significant [17]. Detailed description of constructs used in this thesis as well as description of culturing and assay protocols can be found in the section Materials and methods located in the Appendix.

4 Results

4.1 In silico results

4.1.1 Wild type Ste12 characterization

Simulations of Ste12 wild type model were performed using alpha factor in range from 1 to 100000 molecules. According to the authors [8] this number of molecules corresponds to alpha factor concentrations from 10^{-10} to 10^{-6} M.

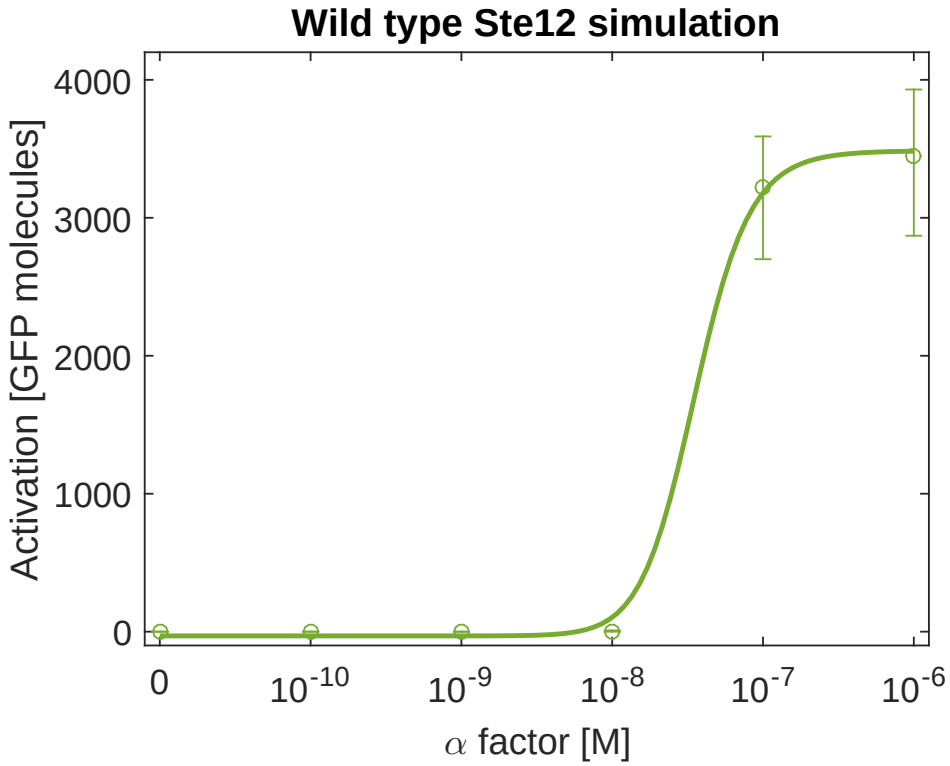


Figure 4.1: FUS1-GFP dose response obtained from simulation

For maximal activation 10^{-6} M was needed. Steady state of the step response was reached after 3500s of the simulation. Obtained data characteristics correspond to experimental results published in [6] (Fig. 2.6A). However data obtained from [7] indicates that actual experimental results are dependent on used method and yeast strain (Fig. 2.6B).

4.1.2 Ste12 variations

Series of simulations for 4 different Ste12 concentrations with alpha factor in range from 1 to 100000 molecules was performed. Simulation time was 4000s which took approximately 25 minutes of CPU time. Since the model is stochastic each particular case was simulated 3 times and mean was calculated. The amount of GFP after 4000s was declared as stable state and a dose response was plotted.

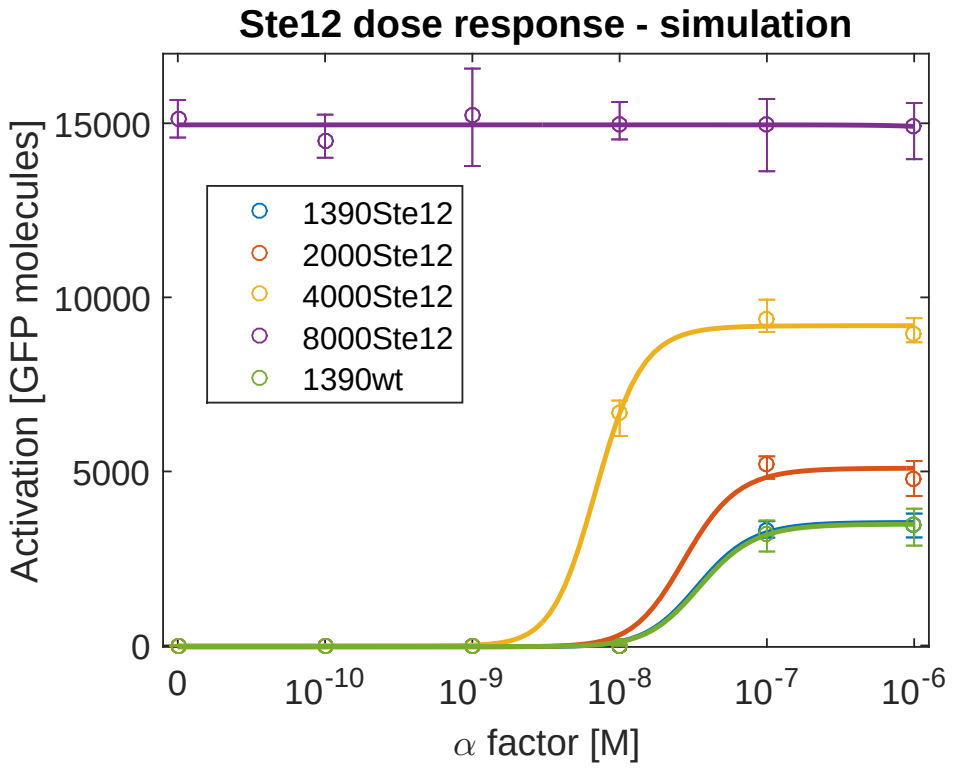


Figure 4.2: Fus1-GFP dose response for different concentrations of Ste12 obtained from simulation

Particular concentration were used for following reasons:

- 1390 is estimated number of molecules of Ste12 in wild type strain used in original model. This concentration was simulated using model where rules enabling self-amplification of Ste12 were included (1390wt). Simulations for this concentration using model without these rules were performed as well (1390Ste12). All other concentration were simulated in model without this positive feedback.

- 2000 and 4000 Ste12 molecules are inactive in absence of pheromone due to their connection with Dig1/Dig2. 2000 is also the maximal

anticipated amount of molecules in wild type strain according to [2].

- 8000 molecules exceed concentration that can be in inactive form in absence of pheromone. This concentration showed no significant differences among all pheromone doses in Ste12 simulations.

Removal of the self-amplification feedback showed no noticeable difference for 1390 molecules. 4000 Ste12 molecules concentration was active at lower concentration (10^{-9}M) and seems to be more sensitive.

4.1.3 PAF

Similarly to Ste12 series of simulations of PAF model was performed using same concentrations and conditions as in Ste12 case.

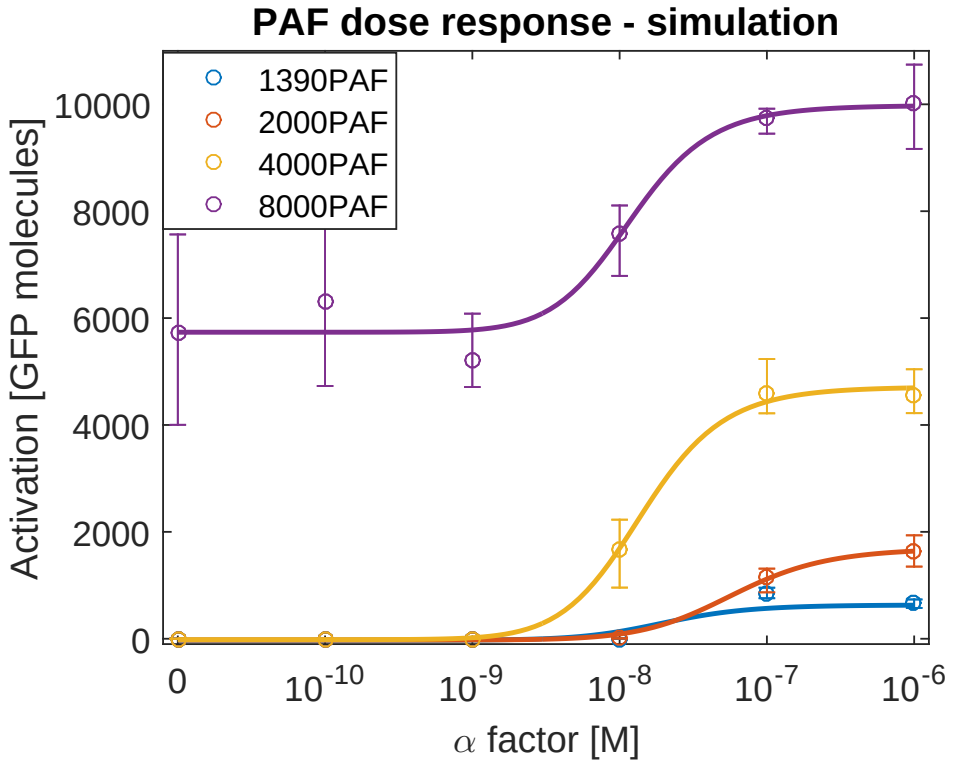


Figure 4.3: GAL1-GFP dose response for different concentrations of PAF obtained from simulation

As expected the dose response curves correspond to the Hill function. Lower concentrations of PAF are inactive in presence of low alpha factor concentrations whereas concentration of 8000 molecules is active even in absence of pheromone. Saturation requires approximately 10^{-7} M of alpha factor. Amplitude varies for different concentrations.

Significant distinction between Ste12 and PAF is higher maximal activation of Ste12 for each used concentration caused by the positive feedback. For instance activation corresponding to approximately 10000 GFP molecules requires 8000 PAF molecules in contrast to only 4000 Ste12 molecules. As mentioned earlier dose response of 8000 Ste12 molecules is basically constant. This indicates that this concentration reached maximal possible activation of FUS1 promoter (15000 GFP molecules). This phenomenon was not observed during 8000 PAF molecules simulations since maximal reached concentration was 10000 molecules.

4.2 In vivo results

4.2.1 Wild type Ste12 characterization

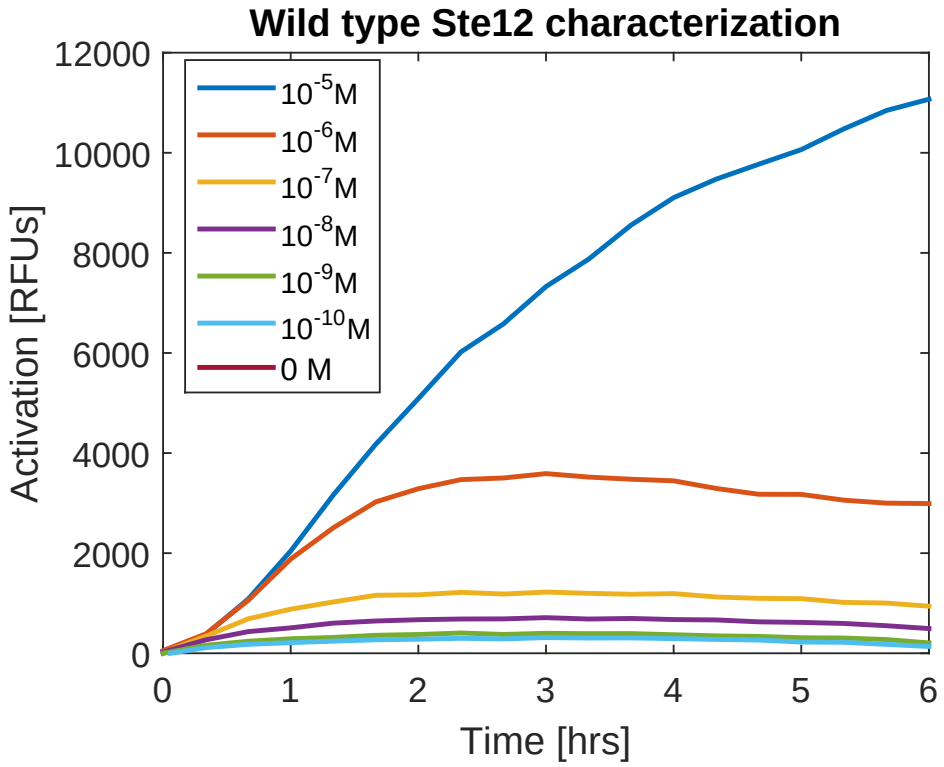


Figure 4.4: Measured step responses of FUS1-GFP reporter characterize activation of wild type Ste12 transcription factor in presence of various alpha factor concentrations

Response of wild type Ste12 protein to alpha factor was measured in 6193 strain transformed with plasmid carrying pFUS1-GFP reporter. Doses of

alpha factor in range from 10^{-10} to 10^{-5} M were used. Minimal input in form of pheromone for activation of wild type Ste12 was 10^{-7}M. There is noticeable delay in fluorescence production caused by transmission of the signal throughout the pheromone response pathway and maturation of the GFP protein. Steady state was reached after 2 hours except for the highest alpha factor concentration. Unfortunately saturation was not achieved, even though graphs obtained from both literature and simulation (Fig. 4.1) indicated that maximal activation should be reached at dose 10^{-6}. It seems that the method of measurement and the used strain may significantly influence the result (see Fig. 2.6). Dose response of wild type Ste12 activated by alpha cells is included in following section.

4.2.2 Ste12 variations

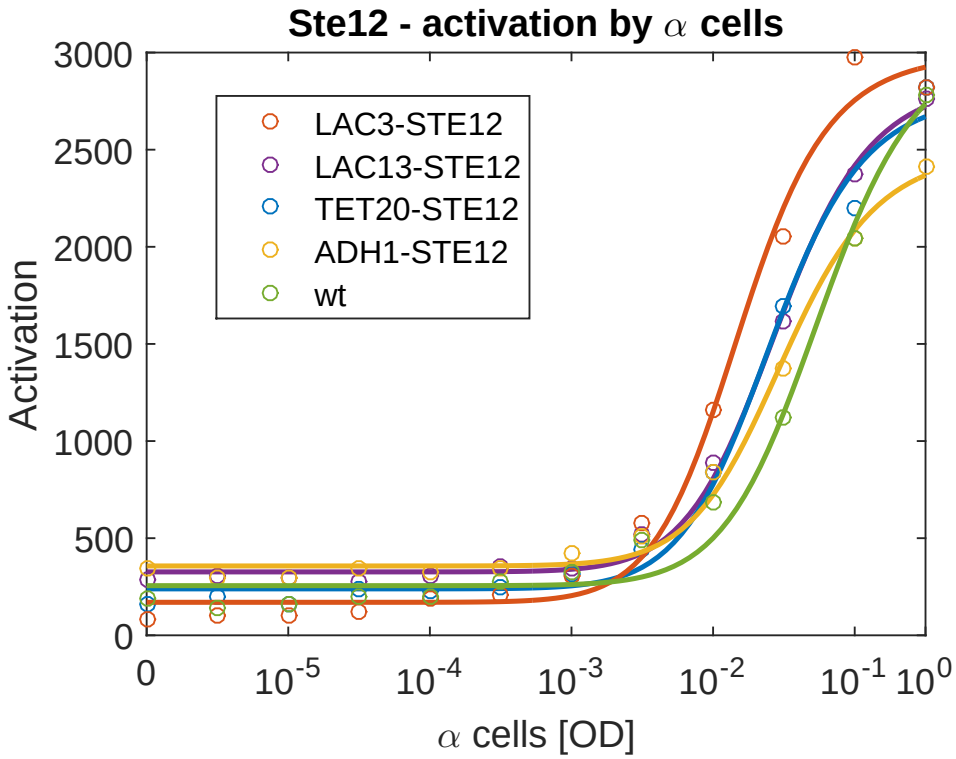


Figure 4.5: Experimental characterization of Ste12 transcription factor regulated by different promoters using activation by alpha cells

Alpha cells of various concentrations were added to a-cells cultures containing Ste12 regulated by several promoters. Higher OD of alpha cells was used for sake of achievement of the saturation. Activation of all cultures was successful. There is small vertical shift among the dose responses

visible at the lowest alpha cell doses. Ste12 regulated by pADH1 showed lowest maximal activation. However results showed only minor differences among all promoters. Research performed by [18] indicates that existence of Ste12/Dig1/Dig2 complex enables creation of large supplies of inactive Ste12 prior pheromone stimulation which allows rapid response after pheromone stimulation. Ste12 molecules beyond maximal amount of this complexes are unstable and undergo quicker degradation. This mechanism may explain data resemblance of all promoters since yeast cell only maintain amount of Ste12 required for appropriate activation level. It is also possible that maximal activation of FUS1-GFP reporter was reached in all cases.

4.2.3 PAF

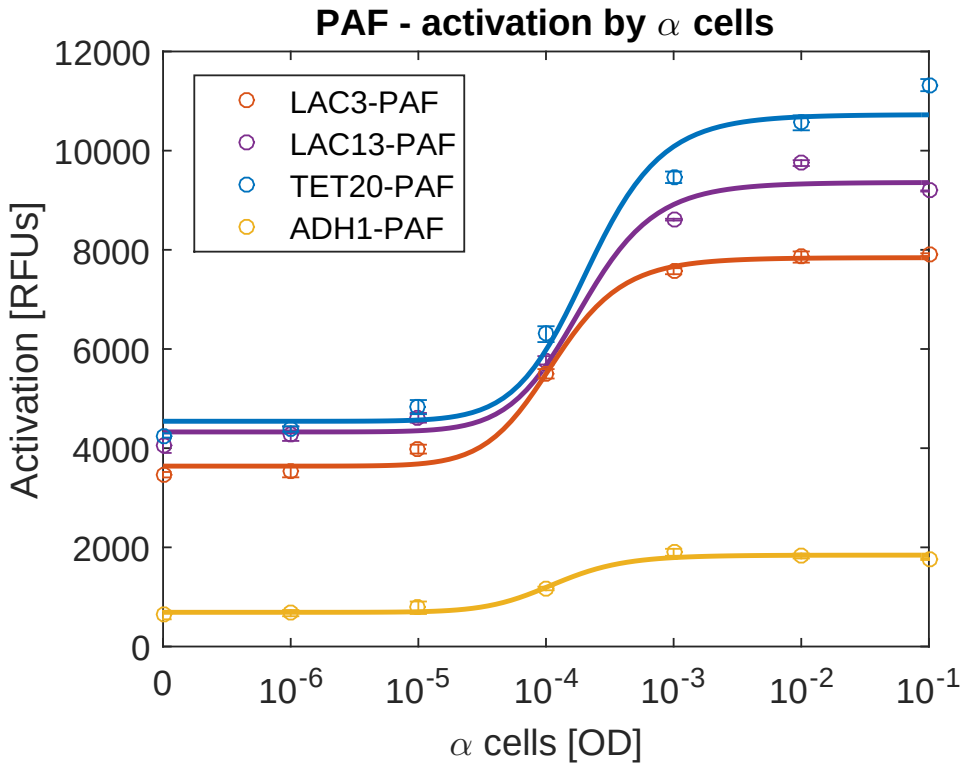


Figure 4.6: Experimental characterization of synthetic PAF transcription factor placed downstream of synthetic promoters using activation by alpha cells

Fig. 4.6 presents GFP dose response curves indicating PAF activation in a-cells after addition of alpha cells. Alpha cells in range from 10^{-6} to 10^{-1} OD were used. Steady state of fluorescence was reached after 6 hours of

the measurement and maximal activation required at least OD 10^{-2}. The amplitude varies depending on different promoters.

As expected dose responses are vertically shifted. The order of three synthetic promoters is same as the order obtained by their characterization, assuming that the higher the fluorescence the higher the PAF concentration. However pADH1 promoter seems to be the weakest one. This is in contrast to expected results.

There is quite high fluorescence even when low alpha cells concentration was added. Model predicted this result for PAF concentrations exceeding maximum number of Dig1/Dig2. This behavior might be also caused by decreased ability of Dig1/Dig2 to bind PAF. This would lead to its constant activation. Also the degradation rate of PAF is not known. As mentioned earlier Ste12 is stabilized by Dig1/Dig2 otherwise degrades quickly. This characteristic does not have to be same for the PAF protein.

Second experiment using alpha factor induction was performed to prove the ability of dose response modification and to verify the results.

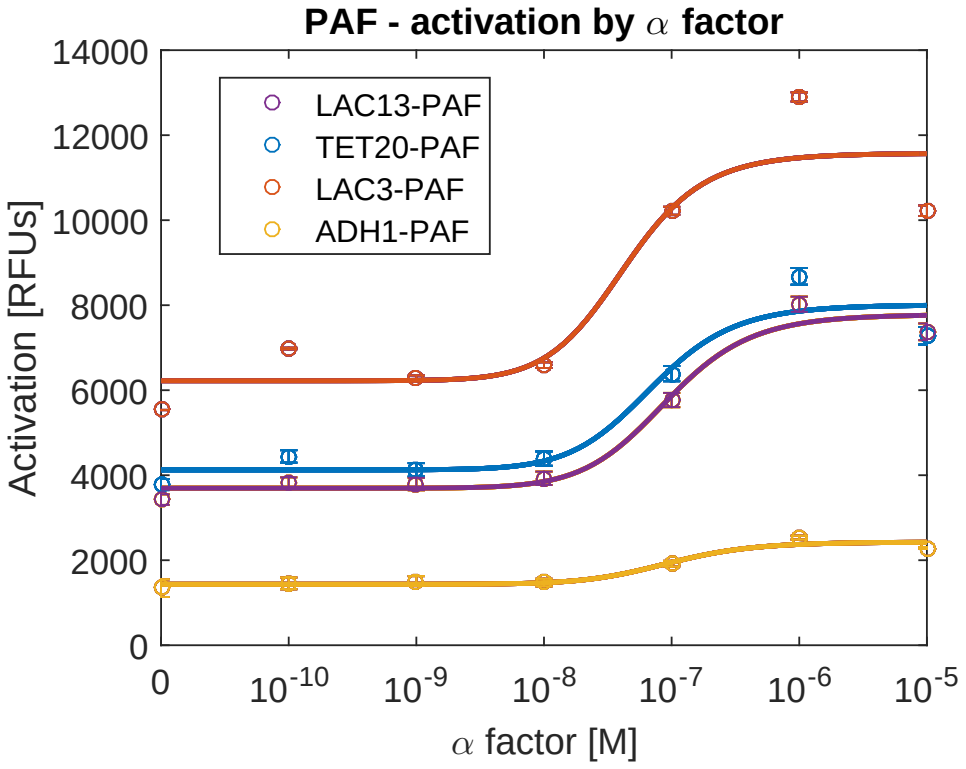


Figure 4.7: Experimental characterization of synthetic PAF transcription factor placed downstream synthetic promoters using alpha factor doses

Alpha pheromone in range from 10^{-10} to 10^{-5} M was used. Steady state

was reached earlier than after activation by alpha cells (after 3 hours) and maximal activation required at least 10^{-6} M of alpha factor.

The possibility of vertical shifting of the dose response was successfully confirmed. The activity of PAF increased at least twice after alpha factor addition.

The order of promoters does not correspond to the previous case. Each repetition of this experiment showed different order of these promoters. As mentioned earlier the difference in strength of the three synthetic promoters is not so significant and therefore their activity might differ. Once again pADH1 promoter seems to be the weakest. This phenomenon has not been explained yet. Literature indicates that it might be necessary to use Gal4 and Gal80 knockout strain when using GAL1 promoter as a reporter. Gal80 represses Gal4 and thus it is possible that it also represses PAF [19].

However since both pADH1-Ste12 and pADH1-PAF seem to cause the lowest production of GFP and this promoter should produce high Ste12 concentrations, it is possible that higher amount of transcription factor does not cause higher activation as expected.

5 Disscusion

This work focused on Ste12 and PAF transcription factors activated by yeast pheromone response pathway. Theoretical research summarized in second chapter brought up important facts necessary for correct understanding of the systems. This knowledge was used during development of mathematical models followed by adaptation of complex stochastic yeast pheromone response model. Simulation results were than experimentally validated.

The synthetic transcription factor PAF was chosen as the alternative of the wild type Ste12 transcription factor due to its ability to activate galactose inducible promoters without activating the mating response. PAF can be used in further researcher as a transmitter of signal detected by yeast pheromone response pathway and can trigger many different responses.

In silico results predicted possible vertical shift of activation dose response caused by different concentrations of the transcription factor. Ste12 enables higher activation then PAF thanks to it positive feedback function. Tuning of the response is therefore possible.

In vivo, wild type Ste12 was characterized. Experiments for comparison of 4 different concentrations of both PAF and Ste12 were performed. Results obtained for PAF confirmed predicted vertical shift. Ste12 experiments showed only slight differences among all concentrations however signs of the vertical shift were observed even during these experiments. These ambiguous results demonstrate that even though the yeast pheromone response pathway is well characterized there are still findings that need to be further examined.

List of Figures

Bibliography

[1] Tereza Puchrová. Modelling and experimental validation of signalling pathways with relevance to homologous mammalian systems. Master's thesis, University of West Bohemia, September 2015.

[2] Lee Bardwell. A walk-through of the yeast mating pheromone response pathway. *Peptides*, 26(2):339–350, February 2005.

[3] Raymond E. Chen and Jeremy Thorner. Systhematic epistasis analysis of the contribution of protein kinase a- and mitogen-activated protein kinase-dependet signaling to nutrient limitation-evoked responses in the yeast *Saccharomyces cervisiae*. *Genetics*, 185:855–870, July 2010.

[4] Haiwei Pi, Cheng-Ting Chien, and Stanley Fields. Transcriptional activation upon pheromone stimulation mediated by a small domain of *Saccharomyces cerevisiae* ste12p. *Molecular and Cellular Biology*, pages 6410–6418, 1997.

[5] Step response. [online], December 2015. [cited 2016-05-3] Available at: `https://en.wikipedia.org/wiki/Step_response`.

[6] Diener C, Schreiber G, Giese W, del Rio G, Schröder A, and Klipp E. Yeast mating and image-based quantification of spatial pattern formation. *PLoS Comput Biol*, 10(6), June 2014.

[7] Mark A. Poritz, Sharon Malmstrom, Marianne K.-H. Kim, and Peter J. Rossmeissl nad Alexander Kamb. Graded mode of transcriptional induction in yeast pheromone signalling revealed by single-cell analysis. *Yeast*, 18:1331–1338, March 2001.

[8] Ryan Suderman and Eric J. Deeds. Machines vs. ensembles: Efective mapk signaling through heterogeneous sets of protein complexes. *PLoS Computational Biology*, 9(10), 2013.

[9] Bionetgen tutorial. [online], 2015. [cited 2016-04-11] Available at: `http://bionetgen.org/index.php/BioNetGen_Tutorial`.

[10] A. M. Smith, W. Xu, Y. Sun, J. R. Faeder, and G. E. Murai. Rulebender:integrated modeling, simulation and visualization for rule-based intracellular biochemistry. *BMC Bioinformatics*, 13(Suppl 8):1–24, June 2012.

[11] Biological half-life. [online], September 2016. [cited 2016-04-18] Available at: `https://en.wikipedia.org/wiki/Biological_half-life`.

[12] Daniel T. Gillespie. A general method for numerically simulating the stochastic time evolution of coupled chemical reactions. *Journal of Computational Physics*, 22:403–434, 1976.

[13] Michael W Sneddon, James R Faeder, and Thierry Emonet. Efficient modeling, simulation and coarse-graining of biological complexity with NFsim. *Nature Methods*, 8(2):177–185, February 2011.

[14] Brian E. Munsky. Stochastic chemical kinetics. Presentation.

[15] Green fluorescent protein. [online], 2016. [cited 2016-04-26] Available at: `https://en.wikipedia.org/wiki/Green_fluorescent_protein`.

[16] Tom Ellis, Xiao Wnag, and James J Collins. Diversity-based, model-guided construction of synthetic gene networks with predicted functions. *Nature Biotechnology*, 27(5), April 2009.

[17] Hynek Kasl, Tereza Puchrová, and Daniel Georgiev. Unpublished data.

[18] John R Houser, Eintou Ford, Michal J Nagiec, Beverly Errede, and Timothy C Elston. Positive roles for negative regulators in the mating response of yeast. *Molecular Systems Biology*, 8(article n. 586), April 2012.

[19] Clontech Laboratories Inc. Yeast protocol handbook, July 2009.

[20] Michael C. Lorenz and Joseph Heitman. Yeast pseudohyphal growth is regulated by GPA2, a G protein α homolog. *The EMBO Journal*, 16(23):7008–7018, June 1997.

A Materials and methods

Yeast strains and methods. Yeast strains used for the assays were MLY215 ($\Delta pde2$ :: $G418\Delta$ $ste12$ :: $leu2$ :: $hisG\Delta$ $leu2$ :: $hisG$ $ura3-52$ $MATa$) [20] in case of Ste12 with synthetic promoters and PAF measurements and 6193 ($FY23$ :: $ura3-52$ $leu2\Delta1$ $trp1\Delta63$ $MATa$) for characterization of wt. Ste12. Yeast strain used for induction was 6194 ($FY23$ :: $ura3-52$ $leu2\Delta1$ $his3\Delta200$ $MAT\alpha$). All constructs were transformed into corresponding strains using High efficient yeast transformation. pRS416 - pGAL1-GFP reporter plasmid was used in PAF assays. pRS416-pFUS1-GFP reporter plasmid was used in Ste12 assays. MLY215 (MATa) (6193 (MATa)) strain containing empty pRSII415 and pRS416 vectors was used as negative control of fluorescence in PAF and synthetic Ste12 assays(wt. Ste12 assay).

Plasmid construction. All constructs were validated on gel and sequenced. Description of construction of plasmids created directly for this thesis follows.

1. **PAF protein.** Complete sequence of PAF gene was design and ordered as a G-block using description in [4]. G-block was ligated into pRSII415 vector using BamHI and SpeI restriction sites. Synthetic promoters pLAC3, pLAC13, pTET20 were design, ordered in form of G-blocks and validated by other members of the laboratory using [16]. PAF digested using HindII, SalI enzymes was ligated into previously created cassette containing pADH1 promoter and asCYC1 terminator in pRSII415 vector. Constructs containing promoters and pADH1-PAF-asCYC1 construct were digested using XhoI and HindIII restriction enzymes and the synthetic promoters were ligated upstream the PAF sequence.

2. **Ste12 with synthetic promoters.** Sequence of Ste12 was amplified from genome through PCR. PCR product containing Ste12 sequence and PAF constructs with promoters and terminator were digested with HindIII, SalI enzymes. Ste12 was ligated within the promoter and the terminator.

Fluorescence intensity measuring assay. Colonies after transformation were streaked on selection plates with appropriate amino acids. Liquid cultures in SD —ura —leu dropout medium were prepared and grown for

24 hours. Cultures were diluted to OD 0.05 in case of cells induction and OD 0.15 in case of alpha factor induction. Diluted cultures were grown for extra 4 hours before the measurement. Cultures induced by alpha cells were diluted to OD 0.1. Cultures of alpha cells used for induction were grown in SD —ura —leu dropout medium for 24 hours, refreshed and grown for extra 4 hours. Alpha cells were then washed with water and 1xTE to remove alpha factor. Fluorescence and OD were measured for 7 hours within 20 minutes intervals using plate reader. Each sample had 2 biological replicated however differences between them were minimal.

Data processing. Fluorescence and OD of blanks were averaged and subtracted from each sample. Ratios fluoresce/OD were calculated for each sample and negative controls and average between replicates was calculated. Corresponding negative control was subtracted from each sample. Dose responses of steady states were plotted at time when all samples had already reached maximal activation.

B List of Abbreviations

BNGL BioNetGen language

DNA deoxyribonucleic acid

GDP guanosin diphosphate

GFP green fluorescent protein

GTP guanosin triphosphate

MAPK mitogen activated protein kinase cascade

NFsim Network-free simulation algorithm

OD optical density

ODE ordinary differential equations

PAF pheromone activated synthetic transcription factor

PCR polymerase chain reaction

PRE pheromone response element

RNA ribonucleic acid

SSA Stochastic Simulation Algorithm

UASGAL upstream activation sequence of galactose inducible promoters

XML Extensible Markup Language

YPRP yeast pheromone response pathway