# MAPPING AND MODELING GENETIC AND MORPHOLOGICAL PHEROMONE RESPONSE IN FISSION YEAST

MASTER THESIS

# PRESENTED TO THE FACULTY OF BIOSCIENCES

AT THE RUPRECHT-KARLS-UNIVERSITÄT HEIDELBERG

by Ábel Vértesy

2012

2 / 120

This thesis was written as a part of the Molecular Biosciences master program at the University of Heidelberg in the period of 07/05/2012 to 31/10/2012 under the internal supervision of Professor Ursula Kummer. The thesis was accomplished externally at the department of Theoretical Biophysics at Humboldt University under the supervision of Professor Edda Klipp.

1 <sup>st</sup> Appraiser:	Professor Ursula Kummer,	Institute of Zoology
2 <sup>nd</sup> Appraiser:	Professor Dirk-Peter Herter	n, Institute for Physical Chemistry

I herewith declare that I wrote this Master thesis independently under supervision and used no other sources and aids than those indicated.

Date

Signature

4 / 120

# TABLE OF CONTENTS

Foreword	7
Acknowledgement	
Biology	9
Evolution and homology	9
Differences to cerevisiae, application	9
Cell cycle	9
Cell type and mating	
Unique features of the pombe pheromone response	
Two faces of the response	
Terminology	
Introduction to the problem	15
The question	15
The challenge	16
The approach	17
Methods	
Methods for pathway reconstruction	
Looking for a formal framework for network reconstruction	
Modeling by ordinary differential equations	22
Ordinary differential equations	22
Parameter estimation	23
Applications of modeling in biology	24
Compartment modeling	26
Results and discussion	27
Data and data generation	27
Data generation	27
Time course data and cell morphology	
Genome wide protein expression data	41
Pathway reconstruction results	45
Starting point	45
Genetic regulation	45
Starvational induction	45
Full picture and re-simplification with a different focus	47
Explaining Spk1 activity and morphology in the same model	49
An explanation for Ras1	50
Designed experiments	51

Approaching the function of Gpa1	54
'Conservation of Spatial Information' concept helps to constrain the possible set	t of 54
Possible Gpa1 effectors	
The 'Conservation of Spatial Information' concept (CSI)	55
Consequences of CSI on possible effectors	55
Model - Mark I	60
Model - Mark II	66
Further development	70
Annotation	71
The Ras1 colocalization model	72
Implementation	74
Results and discussion	75
Models for subsequent activation peaks	77
Model 1: Different activators model	77
Model 2: Self-sensitization model	78
Results and validation trial	78
MAPK models	80
Alternatives in negative regulation	80
Transcriptionally controlled regulators	81
Modeling a general MAPK module with a delayed downregulation	82
Opposite effect of loop-length and distance from point of regulation	83
Gpa1-Rgs1 model	85
The general explanation of Gpa1 behavior	85
Analytic confirmation of the minimum number of states of Gpa1	
The original model	90
Alternative models	91
Results	91
Cell polarization in pheromone signaling	95
The distance measurement hypothesis	98
Shmooing in pombe	98
The hypothesis	99
Support for the hypothesis	99
A plan to select the correct hypothesis	101
Conclusion and Outlook	103
Appendix	106
Homology with cerevisiae and general names	106

Alternative names of pombe proteins	
Sex-specific genes	
Online resources used in the project	
Summary of mutant strains	
Transcriptional regulation	
Subcellular localization of PPW-proteins	
List of Supplementary material	
Bibliography	

## FOREWORD

Systems biology brought –among others- two major advances to modeling and theoretical sciences. One it became a standard that the model itself has to be published along with the paper, and second is that models are highly standardized, easing assessment, exchange and reuse. In my opinion, there is still one major step that needs to be made to maximize the acceptance and therefore influence of scientific modeling. That is providing full datasets used for model construction and model fitting. Reproducibility and reliability suffers from this practice, therefore I would like to foster publishing data along models. Therefore datasets are also provided along this work. Even more, context of data generation or metadata is provided, acknowledging its importance. [1]

This thesis contains references to non-peer reviewed sources, for example to Wikipedia and to the Qiagen GeneGlobe pathway repository. I think there is an important place for non-scientific references in scientific works in certain cases. These sources have the advantage over often very technical and narrowly focused papers, in that they are written in an educative manner, and they are much easier to understand at first glimpse. Nonetheless one should stick to peer-reviewed references if the cited information is not widely established or crucial regarding the conclusion.

30/10/2012

### Acknowledgement

I am grateful to Professor Ursula Kummer for her useful comments and for helping me in the search for a thesis project. I would like to especially thank Professor Edda Klipp the supervision, the great discussions and creating a place like the department of Theoretical Biophysics with a very inspiring and relaxed atmosphere. I am indebted to Dr. Kayoko Tanaka and Emma Kelsall for a truly productive and pleasant collaboration. I am obliged to many of my colleagues who supported me, I am especially grateful to: Dr. Gabriele Schreiber for her help with the experiments, Ivo Maintz for helping so many times with all kind of software; Dr. Matteo Barberis for the nurturing discussions about the Gpa1-Rgs1 concept; Wolgang Giese for his ideas about spatial modeling and cell polarization models; Dr. Ercan Engin Kuruoglu for the introduction to Bayesian network reconstruction methods; Max Flöttman and Dr. Samuel Drulhe for a series of discussions; Professor Alejandro Colman-Lerner and Dr. Marcus Krantz for their useful comments and for sharing their knowledge about signaling in budding yeast; Professor Lars Kaderali and Matthias Reis for his useful hints on the analysis of mRNA and Protein data; Professor Alexander Hoffman for his ideas about terminology and visualization; Dr. François Guillaud, Claudia Beck and Katja Tummler for helping with 3D visualization; Matthias Bock for reviewing the Manuscript. Last but not least, I am grateful to my parents whose care brought me to the end of my studies and to Alicia Torrea who supported me during this work.

# BIOLOGY

#### Pombe and the pheromone response system

#### **EVOLUTION AND HOMOLOGY**

Schizosaccharomyces pombe, or fission yeast is an Ascomycota fungi and is about 300 – 400 million years ago separated from cerevisiae. [2] Nevertheless the remoteness in evolution they live in very similar ecological niches (habitat) Both obtained their names from beer; cerevisiae from Latin word, pombe from the East African word for millet beer. The similarity in habitat though the different evolution yielded many analogies between them: similar attributes with different origin. This makes comparisons with cerevisiae limitedly useful. Naming of proteins is sometimes even misleading. (See: *Homology with cerevisiae and general names*) Nevertheless the, both pathways show similarities: [3].

#### DIFFERENCES TO CEREVISIAE, APPLICATION

Pombe is a genetically stable organism, one of the most common organisms of molecular biology. Its genome was sequenced as early as 2002. [4] While pombe is a bit more difficult to grow than cerevisiae, is in some feature more similar to mammals. An early review on differences: [5] also see: an info graphic from 'The Scientist'<sup>1</sup>. In short:

- 1. Pombe and metazoan divide by fission, not by budding.
- 2. Pombe has larger chromosomes than cerevisiae, also they condensate more during cell division: can be stained and visualized.
- 3. Pombe has around more than 5000 introns in 40% of the genes, whereas cerevisiae has approximately 250 introns in 5% of its genes.
- 4. Pombe has more complex centromeres than pombe.

Pombe is the most commonly used for research of cell cycle, DNA damage, RNA processing.

#### Cell cycle

Cell cycle of Pombe is dominated by a long G2-phase and a major G2 -> M main transition checkpoint, contrasting cerevisiae which mainly resides G1, and where the most important control is G1 -> S transition. Interestingly this difference in cell phase evens out the gene copy number: the diploid living cerevisiae resides in the phase

<sup>1</sup> http://images.the-

scientist.com/content/figures/images/yr2003/pdfs/yeast\_supp\_0306xx.pdf

before synthesis, whereas the haploid stable pombe exists in for most of its time in the phase after the S-phase. This is also a nice example of analogy: same attribute emerges from different evolutionary origins.



FIGURE 1: POMBE CELL CYCLE SHOWING THE RELATIVE LENGTH OF PHASES. DAUGHTER CELLS OFTEN STAY TOGETHER AFTER DIVISION. CELLS COMMIT FOR CONJUGATION UPON STARVATION BY ARRESTING CELL CYCLE AT G1 BY STARVATION. SOURCE: K. TANAKA, MODIFIED.

#### **Cell type and mating**

Fission yeast has two mating types, plus and minus; cells correspondingly secrete Pand M-factor. [6–8] Aside from a few interesting differences they are very similar both in behavior and pathway components. One of the main differences regards the degradation of pheromone. Sxa1 of P-cells is induced by starvation, whereas Sxa2 of Mcells require pheromone stimulation. [9] Sxa1 is possibly more than a pheromone protease, since it localizes to the ER so differently to Sxa2 ([10] and **Subcellular** localization of PPW-proteins). It was also reported that  $sxa1\Delta$  strains have a general defect in proteolytic capacity. [11], and chapter 26, p. 400 in [12] Sxa1 is expressed in both cells [9], but it is involved in the degradation of M-factor only. [11] Opposing to cerevisiae, pombe mating is only a response to nitrogen starvation. In the model I describe M-cells (responding to P-factor), therefore the naming follows M-cells in the sex-specific genes and proteins. The full list of sex specific proteins is provided in the appendix: **Sex-specific genes**.

30/10/2012

#### Starvation

Fission yeast predominantly grows in haploid state; diploid state is a transient state in response to starvation. Cells conjugate and form diploid zygote. After a single meiotic division, they form four haploid spores, which develop to haploid cells if condition improves. See: [13], chapter 19-20 of [12] and Wikipedia. The sensing of starvation is not clear, but it certainly results in a decrease of cyclic-AMP levels, which indirectly activate Ste11, the master transcription factor of pheromone response and mating. [14], chapter 18.5 of [12] Possible pathway elements are the Stn1 7TM<sup>2</sup> receptor sensing the presence of nitrogen, Gpa2 alpha component of a trimeric G-protein and Cyr1, a cAMP synthase. [15,16] nitrogen starvation is also known to activate the Sty1-Atf1-Pcr1 Stress MAPK pathway, again resulting in the activation of Ste11. [17], [12] Glucose starvation is not necessary, but it facilitates efficient mating. chapter 18.5 of [12] Nitrogen starvation also activates mTOR kinase, which accelerates the the G2 -> M transitions, cells finally arrest at G1. [18] Cell cycle can be arrested by N-starvation [18] or by pheromone sensing. [19–21]

#### Pheromone sensing, mating, meiosis and sporulation

Starvation induces the transcription of pheromone and a multitude of pathway components (see: *Transcriptional regulation* appendix), thereby prepares the cell for pheromone signaling. [22,23] and [12] Upon pheromone sensing, the genetic program [24] and cell morphology is changed. If pair of opposite sex cells are close enough, they recognize to each other and grow shmoos. Subsequently cells fuse and finally form a diploid zygote. This zygote has now both sex-specific mat1 loci, therefore can express both mat1-Pc and mat1-Mc. [25,26]

These form a heterodimer transcription factor, and induce the downstream Mat1-Pi and Mat1-Mi loci, again forming heterodimeric transcription factors. This induces Mei3, which inhibits Pat1 kinase that ends Mei2 inhibition. [27] Mei2 is the conductor of the meiotic transcriptional program. [22,23] Cells then commit to meiosis and subsequent sporulation, and 4 haploid spores are formed within the cell.

<sup>&</sup>lt;sup>2</sup> 7TM: Seven transmembrane receptor.

#### UNIQUE FEATURES OF THE POMBE PHEROMONE RESPONSE

The cerevisiae pheromone pathway is a much better described and already mathematically modeled system.[28]. Both pathways prepare the cell for mating, both sense by G-protein coupled receptors (GPCR), utilize trimeric G-proteins, and contain a MAP-kinase cascade. Despite these shared properties, there are many discrepancies among the pathways. [29,30] The pombe PPW does not contain any scaffold proteins, opposed to cerevisiae to our current knowledge. Even if scaffolds exist, it is probably not essential, as no sterility phenotype is identified in knockout screens. It is theoretically possible that some of the enzymes act as scaffold; however it is unlikely because their relatively small molecular weight. Cerevisiae pheromone response does *not* require a previous induction of the starvation-pathway. The pathways follow a completely different time scale: 6-12 hours vs. 10-20 minutes in budding yeast. Often, proteins of similar role have rather different amino acid sequences.

#### TWO FACES OF THE RESPONSE

#### A global genetic change and a local alternation of morphology

What is commonly understood under the 'pheromone pathway' is typically the transcriptional response. It neither covers the full pheromone response, nor a functionally stand-alone pathway. It is widely known that there are two main reactions to pheromone treatment. First, they upregulate several genes whereas possibly downregulate a few others through the activation of Ste11 transcription factor. [14,31] Second, cells grow a shmoo by reorganizing their actin cytoskeleton and the cell wall. It is also established that these responses utilize common pathway elements; nevertheless surprisingly these two responses were never treated in a common framework. In this work 'pheromone pathway' always refer to signaling in both responses. Nevertheless, I am aware that the pheromone pathway is tightly linked to other pathways. Many other pathways converge on Ste11, some activate the Cdc42-Scd1-2-Shk1 complex. [18,32,33];

#### TERMINOLOGY

A common trunk activates both genetic- and morphology-branches.

The full pheromone response (Pheromone PathWay, PPW) is a bifurcating pathway. Although many, here ignored crosstalks connect this pathway to others, the pathway can be split up into three distinct parts, according to their *function* and the *type of information* transmitted. A new terminology was necessary to refer to these characteristic different parts explicitly.

#### Common trunk (COT)

The most upstream components of the pheromone response are completely shared between the morphological- and the transcriptional response, therefore these must transmit the rough temporal signal of changes in expression, and a more precise spatio-temporal signal. This pathway consists only of membrane associated components: the pheromone receptor, **Mam2** [34], G-protein alpha subunit,**Gpa1** [35] and **Ras1** [36]. Activators of the *common trunk* are the cytosolic Ras-GDP->GTP exchange factors **Ste6** [37,38] and **Efc25** [39]; but they may be recruited to the membrane by Gpa1. Efc25 is not regulated transcriptionally by pheromone response<sup>3</sup>, therefore it is not incorporated in the model as variable; it is represented in a parameter. Known negative regulators of *COT* are **Rgs1** and **Gap1**. [40,41] Gap1 is constantly expressed too therefore it is part of corresponding deactivation parameter.

The reason why it is stated that the *transcriptional branch*<sup>4</sup> needs a less precise signal than *morphology branch* is highlighted by the *Byr1.DD* strain (see: *Byr1.DD strain*). Here the genetic program is solely induced by starvation; therefore it is without respect to the individual cells state. Still these cells find their partner correctly. On the other hand, strains which lack any negative regulator in the *common trunk* (*gap1* $\Delta$ , *Ras.Val*, *sxa2* $\Delta$ ) grow shmoos not reaching any partner, although genetic activity in these is relatively similar to WT (way more similar than *Byr1.DD*). See further details in: *The distance measurement hypothesis*.

#### Transcriptional branch (TRAB)

The *common trunk* feeds in is the *transcriptional branch*. This MAPK-cascade based pathway is responsible to expresses pathway components and regulators required for conjugation and mating. In literature the *common trunk* and the *transcriptional branch* together is often regarded as the 'pheromone pathway'. This subjective selection of components, especially under this name is very misleading.

<sup>&</sup>lt;sup>3</sup> See in: http://www.bahlerlab.info/cgi-bin/SPGE/geexview

<sup>&</sup>lt;sup>4</sup> See next sections

This part contains the MAPK kinase cascade, where **Byr2** phosphorylates **Byr1** twice, which in turn double phosphorylates **Spk1**, which then phosphorylates the high mobility group (HMG) transcription factor **Ste11** on two threonine residues. [42,43] The regulation of Ste11 is much more complex though, barely understood and debate is not yet settled. [44–49] It is also known that both Spk1 and Ste11 shuttle between cytoplasm and nucleus [50], but again the conditions and concrete molecular mechanism remained unclear so far. **Pat1**, a Ser/Thr protein kinase seems also important in deactivating Ste11. [27,46] One of the major steps towards the full understanding of the pheromone response should be the clarification of Ste11 and Spk1 regulation and localization. In this project only the activation by Spk1 is considered.

Other players are also suggested to participate in the transcriptional response. **Ste4**, a leucine zipper protein was shown to bind Byr2 and *Ste4* $\Delta$  mutants suggested it as a pathway activator. [51] By the same group **Shk1** kinase was suggested as a activator of Byr2. [52] During this project, we have proven *in vivo* that the complex containing Shk1 is necessary for pathway activity. (See: *scd1* $\Delta$  in *Unquantified western blots*) This was a huge leap forward: it allowed incorporating additional data, interpreting unexplained behavior, predicting and proving links.

#### Morphology branch (MOB)

There is a second sub-pathway downstream of the *common trunk* is the *morphology branch*. This part is responsible for signaling the correct **direction** and **distance** of the partner chosen for mating. In this project only the first step of this pathway is considered, that is the local recruitment and possible activation of the Scd1-Scd2-Cdc42-Shk1 protein complex. This complex is responsible to recruit further proteins involved in the reconstruction of the actin-cytoskeleton. It seems that Scd2 acts as a scaffold [53,54], where Scd1 activates Cdc42 [55,56], which in turn activates Shk1[57].

Since only Shk1 has at least five known effectors<sup>5</sup>, further description of this pathway would be a project on its own, so it is not discussed here. How much the considered components only **transmit** the signal, or how much they possibly **process** it, is discussed in the *Cell polarization in pheromone signaling* section.

<sup>&</sup>lt;sup>5</sup> Qiagen GeneGlobe: PAK Signaling in Fission Yeast

### INTRODUCTION TO THE PROBLEM

#### THE QUESTION

The original objective in this project was to build a quantitative model on the pathway focusing on the effect of Ras.Val hypersensitive mutation. Quantitative modeling was intended to help experimental work, and to give a mechanistic insight into pathway dynamics.

#### The organism

Schizosaccharomyces pombe is one a common object of molecular biology. It is relatively simple compared to mammalian cells, but still eukaryotic. In addition to this, its genetic stability makes it a suitable research organism; it is also a generally well described system, making it suitable for modeling. The pombe pheromone response shows remarkable similarities to mammalian growth factor responses. Both pathways trigger functions of reproduction and are built up from a G-protein coupled receptor (GPCR), Ras1-homologues and a MAPK-cascade that regulates transcription.

#### **Cancer and diseases**

The importance of GPCR-Ras pathway family is enormous. Around every fourth marketed drug targets GPCR [58]; and the human Ras1 homologue contains hypersensitive mutation [59] in approximately in one third of all human cancers. [60,61]

#### Design principles

Beside the concrete relation to human disease, the pombe pheromone pathway is interesting in itself. As pombe is a well described and simple eukaryote, it is an ideal tool to identify design principles and conserved molecular schemes. Understanding concept of the pombe pheromone pathway could help understanding other GPCR-MAPK systems.

# THE CHALLENGE *No canonical pathway on hand.*

Shortly after the start of the project it became clear that no canonical pathway scheme exists. The diagram presented was clearly insufficient to describe the data and a set of interactions were unclear. (See: Figure 2) Construction of a consistent and fully explanatory pathway diagram turned out to be the most challenging task during the project. It included limited mathematical modeling, but thorough literature research and logical argumentation. This is detailed in the *Pathway reconstruction results* section. In a sense it was a hypothesis generating job, which suggestions we then tried to underpin as much as possible. My part in that was designing decisive experiments. The literature sources are fragmented, sometimes even self contradicting. Resolving the contradictions and the maintaining consistency with all demonstrated facts was a hard task, which required a series of intense discussion with the experimental collaborators.

#### Unknown regulators

The known regulators were clearly insufficient to describe the data. Considered negative regulators were either shown not to affect the downregulation of the signal, or stands under no transcriptional or post-transcriptional control. Consequently they could not produce the delayed downregulation required to allow for an overshoot. Therefore the extension of the pathway was necessary.



FIGURE 2: FIGURE REPRESENTING THE INITIAL NOTION OF THE PATHWAY. THIS VERY SIMPLE SCHEME SHOWED THE SYSTEM AS AN IDEAL CANDIDATE FOR MODELING.

#### No clear input signal

The 'least artificial' approach of the collaborators posed quite a big challenge. Instead of using known amount of pheromone for pathway stimulation, the wash-away of nitrogen is the only "quantified" input to the system. Upon starvation cells start to secrete pheromone that stimulates opposite sex cells. See: *Biology* section. Nitrogen-starvation has a separate pathway that converges either on Ste11 and/or on some other components of the *transcriptional branch*. How much pheromone is induced with what dynamics, remains unknown: we do not treat cells, rather we let them treat themselves with pheromone in an unknown way. This of course makes clear why do we have big variance among biological replicates.

#### Scarce data

To reconstruct a pathway and build quantitative model, time course data on a single protein (Spk1) was clearly insufficient. New type of information had to be incorporated and more data had to be generated. Online databases were searched to acquire further data for the system, and a high-throughput study delivered steady state protein data. [62] *in press* The extension of pathway towards the morphological response allowed making use of cell morphology images. Since there was a working pipeline for measuring Spk1 activity in genetic mutants, this was the main information source. Importantly, data generation was divided for different purposes. Qualitative, roughly sampled westerns on possibly many mutants were done for pathway reconstruction and a few westerns with higher temporal resolution in multiple biological replicates were performed for modeling.

#### THE APPROACH

To face the mentioned problems, different answers were developed. Firstly, the profile of the project was changed from a clear modeling project to network reconstruction. Secondly, various and heterogeneous data sources were integrated in order to consolidate the knowledge about the pathway. In early stages, a series of small models were developed on well defined sub-questions, instead of aiming to model the full pathway. These included the pathway induction (*Models for subsequent activation peaks*), the role of localization in signaling (*The Ras1 colocalization model*) and the Gpa1-Rgs1 interactions (*Gpa1-Rgs1 model*).

## Methods

#### METHODS FOR PATHWAY RECONSTRUCTION

After identifying the problems stated in '*The challenge*' section, I started to fill the gaps in the knowledge about the pathway. Construction of a consistent and fully explanatory pathway structure turned out to be the biggest task of the project. It included very limited mathematical modeling, but thorough literature research and logical argumentation. It is detailed in section: *Pathway reconstruction results*. In a sense it was a hypothesis generation job, which we then tried to underpin or falsify as far as possible. Since the literature source is fragmented and sometimes self contradicting, the resolution of contradictions, and the consistency with all proven facts was a very hard task that required periods of intense discussions with the experimental collaborators.

#### LOOKING FOR A FORMAL FRAMEWORK FOR NETWORK RECONSTRUCTION

The question was how to reconstruct a signaling network in a systematic way. I was looking for a formal framework for network inference (or reverse engineering) instead of the manual reconstruction. Network inference is partially a data fusion problem; I tried to find a formal way to incorporate many different sources of information to find out relations among components. On the other hand it was also a modeling framework problem, since modeling can decide among alternative hypotheses. Simulation can therefore be a useful tool for network reconstruction. Therefore I was looking for both different modeling frameworks to choose the right level of preciseness for the available data.

#### **Bayesian methods**

Bayesian algorithms are popular for network reconstruction. New developments in the field allow overcoming the limitation that classical Bayesian networks cannot handle loops, which are the most ubiquitous control motifs in biology. [76], [77] However, applications for real biological problems were disappointing. In one successful example, in an extremely data-rich setup, Bayesian network reconstruction yielded only partial reconstruction of an already known pathway. [78] Similar results turned out from our own experiences. We participated in the DREAM 7 network reconstruction challenge, and were serially faced by false predictions made by the GRENITS dynamic Bayesian network reconstruction algorithm, whereas we found most of the links by manual inspection of the data.

#### **Boolean networks**

A very simplistic model definition is to encode the state space of variables binary and describe their relation with Boolean operators. This framework is generally well suited for qualitative description of dynamics in large networks. The problem here was the compatibility with our data. The main information sources are semi-quantitative time resolved protein-activity and -concentration measurements. The Boolean framework however loses the amount of input information, by setting measurements either to one or zero. This means, we actually decrease the amount of information that can be used to reconstruct the network.

#### Rule based model generation

One very appealing, and more flexible framework though would be to use rule- and based model definition (generation). The concept behind rule based model generation is to define a model by stating elementary bits of information either as rules or contingencies (constraints). [79] By splitting up the information into these two categories, many assumptions become clearly stated. This is not the case for model definition in general, when one comes up with a subjectively achieved model structure, and focuses on parameters.

In signaling complex formation and multi-site post transcriptional modifications are possible. These give rise to combinatorial complexity: a protein complex of four members may assemble 4! = 24 ways, or a protein with 4 distinct phosphorylation sites can be phosphorylated in 24 sequences. To describe all ODE's in this example is very laborious. Instead the definition of possible state transitions is much simpler: phosphorylation of 4 sites that cover 24 concrete possible reactions. ODE's (or an SBML model) can be automatically generated, which is very handy if the model is complex. Not only model definition, but also ODE-simulations may become unfeasible in highly complex models. This happens if the number of states is in similar magnitude as the number of variables. For this case, agent based modeling is the method of choice. The rule-based model definition software RXNcon (See in section *Software*) is also capable to export the model for the agent-based simulator NFSime through BioNetGen format. In the following some basic concepts are given about rule-based modeling.

#### ÁBEL VÉRTESY

#### Rules for state transitions

Rules describe so called "elementary reactions", also termed "decontextualized reactions" or "state transitions". Formatting state transition rules is compact way to define the full set of combinatorially possible reactions. A state transition rule can be formalized if the two molecular state of an e.g. protein is known. (ERK and phospho-ERK). An example for a state transition rule: A must can form a complex with B and C (but we do not state in which order).

#### *Contingencies on state transitions*

Constraints or contingencies are restrictions on the reaction space. As these are stated separately, one must list all constraints that one knows or one assumes. In this way the modeler explicitly states all assumptions and bits of knowledge regarding what order can (or cannot) a state transitions happen. An example for a contingency: A must interact with C, before interacting with B.

#### Software

The ReactionCon project (RXNcon) offered a server based application for rule-based model generation with a wealth of formal visualization methods. [80] Although formal graphical representations are very explicit therefore a good practice to use them to exchange ideas among modelers, they are often overly complex, distracting the attention from the main structure. They are also very limited and strictly regulated in the use of colors. See: in Figure 14 also in [80] and [81]. For first-glimpse understanding, manually created graphs can be more useful. Since the project was based on the interaction with an experimental group, custom visualization was chosen.





#### Conclusion

Network reconstruction tasks in general need a formal framework that clearly handles the: bits of proven knowledge (a), each assumption (b) and generates all possible hypotheses (c). With such a framework the whole development process could be stored in a standardized way, making the important information of negative results (e.g. refused hypotheses) accessible for others.

Rule based modeling could only offer a half-solution compared to manual network generation, when it comes to be systematic about assumptions and possible hypotheses. Although it is explicit regarding possible concrete reactions (by stating state transitions and all constrains), it still not clearly dissect assumptions and information about the state transitions itself. The most common problem I faced is that functional interaction is known from genetic experiments, but the direct physical interactions or the concrete molecular mechanism is unknown. Consequently state transitions are also unknown. For network reconstruction I would have needed a framework that steps further back than rule-based model definition, and generates possible state transitions. These state transitions denote the distinct hypotheses about the pathway. Since the visualization part would not be used anyway, the half-explicit rule-based definition was not applied either.

The pathway was reconstructed in a manual manner. I tried to enhance reproducibility and reliability by systematically annotating possibly most bits of information (see model file and annotation in supplement); and possible alternatives hypotheses are discussed in text and figures. The manual way turned out to be successful in proposing a limited set of hypotheses testing experiments that were carried out. The trade off was that not all possible hypotheses are considered. Nevertheless, the approach can be considered systems biology since what was finally done is data fusion of various *omics* sources: localization, viability, gene expression, steady state protein counts, interactions from various databases. Merely the way of data fusion was not done in a formal way. Finally, I was successful in that many experiments validated the predictions, whereas some others falsified them, thereby limiting the set of possible hypotheses.

Another conclusion was that strict formal methods can only use formalized input. Most of the knowledge lies in papers. It is often unquantified and unformatted. There are only very limited and complicated methods to quantify literature data, e.g. co-occurrence of terms by text mining. [82] Therefore if one uses a strict formal framework, one probably has to put away most of the available knowledge, unless she/he finds a way to turn it into numeric information.

#### MODELING BY ORDINARY DIFFERENTIAL EQUATIONS

#### **ORDINARY DIFFERENTIAL EQUATIONS**

Since it there are plenty of good introductions for differential equations and nonlinear dynamics, I would avoid creating yet another one. I read and can recommend the concerning chapters in the following books: Alexander Panfilow's Reader on differential equation, [83] and [84]. For an extremely summarized introduction please read on.

A *differential equation* for an unknown function relates the values of the function itself to its *derivative* <sup>6</sup>. The derivative of a function describes the rate of change of that function with respect to the change of an *independent variable*, most typically time. Most often in reality things depend on each other, meaning the change of one entity affects the others. This means, the *dependent variables* in the differential equations describing these things also depend on each other: that yields a set of *coupled equations*. When to equations are coupled, i.e., the solution of one depends on the solution of the other, we call them a *system of differential equations*. Since only the simplest differential equations or systems (of differential equations) can be solved analytically, in a more difficult case one could either restrain to qualitative analysis by concepts of *dynamical systems theory* 

<sup>&</sup>lt;sup>6</sup> From Wikipedia: Differential Equations; modified.

or use *numerical methods* to approximate the solution by extensive computation. Numerical methods are widely applied in many fields, where *quantitative* description and prediction is needed. This work also bases on numerical methods to solve the system of *ordinary differential equations* that describe the pombe pheromone response. An ordinary differential equation describes temporal change of a variable, whereas *partial differential equations* describe spatio-temporal behavior. The spatial side of the dynamics is not treated in this case, the modeled system is assumed to be well mixed, meaning that diffusion speed is not limiting the reaction rates. Rough spatial representation though is introduced in form of compartments. (See: Compartment *modeling*). An ordinary differential equation is a *deterministic* representation of the object, meaning from a given initial state always the same final state is achieved, and the variable follow the same trajectory. One could treat the effect of extrinsic and intrinsic noise with methods like *stochastic differential equations* or stochastic interpretation of the ODE's. As a first description of the pheromone response based on population average data, it is appropriate to use ODE's as they are simpler than the other methods. Other modeling approaches are also mentioned in the *Pathway reconstruction results* section, as modeling is also applied for pathway reconstruction.

#### PARAMETER ESTIMATION

Firstly one creates a system of ordinary differential equations (in further the *model*) that describes the target of observation (in further the *system*), and incorporates data about the initial state of variables. After the model is simulated, it typically does not describe the observations. The solution however also depends on unknown parameters. One could than ask the question: how should I change the parameters, to describe the observations? This problem is often termed an *optimization problem*, as the difference between simulation and observation is to be minimized. Beside random trials, intelligent optimization algorithms exist to find the optimal parameter set. Derivative based methods make use of the local shape of the objective function that describes the difference between the simulation and the data. In high dimensional (i.e. high number of variables) systems, the objective function often has multiple *local* minima. To find the global minimum, one could use global optimization algorithms. *Particle swarm* is one of such, and it was most widely used throughout this project. [85] Tasks were often performed in a multiplex setup, where 10-50 global optimizations were started in parallel from random starting points. In this way both the random

starting point and the global algorithm facilitate to scan the whole range of parameter space, however it is still not guaranteed that the optimal solution is finally found. Another term used in the text for parameter estimation is *model fitting*, grasping the other face of this process. I used Copasi v35 and v39 [86] for modeling, Microsoft Excel 2007 and R v2.14 for data analysis and Inkscape 0.48 and MATLAB for visualization.

#### APPLICATIONS OF MODELING IN BIOLOGY

Although pombe is a top conducted organism, no dynamic (nor spatial) model publicly available describes the pheromone response. Therefore the results of the thesis can serve as prior for starting systems biology research of S. pombe.

#### Mechanistic insight

One typical task of modeling is to highlight the governing laws of the usually obscure and complex biological system. Simple models with variables and links not corresponding one-to-one to biological entities like proteins are common in classical mathematical biology. Such models are useful to highlight the simplest underlying laws. In a sense, such a top-down model in this project is *The Ras1 colocalization model* that describes Spk1 activity by minimal number of variables. Variables in the model stand for "pathway-chunks" and each are representing a set of subsequent proteins in an activator cascade.

#### Pathway reconstruction

Modeling was also applied for pathway reconstruction. In this, alternative simple hypotheses were distilled in to equations, and simulations and parameter estimation were used to prove whether the hypothetic model structure can describe the data. Such an approach was the '*Models for subsequent activation peaks*', when modeling helped to decide between two possible pathway hypotheses.

#### Testing of complex hypotheses and quantitative description

The pathway reconstruction yielded a pathway structure, which was qualitatively consistent with genetic experiments. Although numerous experiments were performed, not all of the 54 reactions were confirmed by us or by others. Some links however still to be assumed in order to achieve a connected graph, or to account for a described knockout phenotype. One may have a guess about how this pathway could behave, but no one is really capable to predict how 53 interacting proteins would behave together.

A way to test such a complex hypothesis for consistency (i.e. that it can behave as it is observed), is to translate it to the explicit language of mathematics. Once these biochemical reactions are written together in a model, it can be tested, whether they can describe the data at all. This is not an easy step, since a hypothesis formulated in text or thought often has hidden assumptions that come up only if one writes them down explicitly. Even if these are clarified, the system may not work as expected before.

The last part of my work was to create a model of differential equations that quantitatively describes the transcriptional pheromone response of fission yeast. This means to find a model structure that can produce a behavior similar to observed, than find a parameter set by parameter estimation, with that the model precisely describes the data. Since each measurement is a composite of the true value and a noise component, precise does not mean exact description. Finding a parameter set, where the model precisely describes the data in a 40 dimensional space is a difficult and time consuming task. Also one could never be entirely sure, if the model is indeed incapable of describing the data. It is well possible that simply the corresponding parameter set is not found.

#### Predicting system behavior

After one arrived at a model that well describes the data, principally there are still infinitely many possible models that can describe the observations just as good. So the modeler is not sure, whether he has the right model. The knowledge about molecular interaction networks constrains the possible alternatives, and also give some guarantee in a sense, that the model is not completely wrong. However one would like to have some support that a good model is finally found. What good means is a difficult question. A model cannot be the one right model for the system, since every model is a simplified description of reality, so for every model exists a limit, thereafter these simplifications become important, so the basic assumptions are not fulfilled. Still, within some limits, using wisely chosen assumptions, a model can be good if it is predictive. A predictive model means that given the model fitted to some dataset; it can guess what will happen in subsequent dataset. This could simply be a longer time course, or measurement of other proteins, etc. In the case of this project the final model predicted genetic variants.

#### **COMPARTMENT MODELING**

Rough spatial representation in ODE-based models is possible by introducing compartments. With compartments one does not assume one well mixed "reactor" for the whole system, rather a set of coupled, well mixed reactors. Typically cellular organelles or membrane vs. cytosol are represented by compartments. In the Ras1 colocalization model I show that the localization of proteins can be as important as a biochemical activation. In this simple model the stimulus mediates translocation and concentration of two species; this catalyzes the reaction between them. With splitting up e.g. the membrane into a series of compartments one can approach spatial resolution; this is a strongly discretized approach points in the direction of PDE's. Such an approach is proposed in the *Cell polarization in pheromone signaling* section; however it was never implemented in a model.

## **RESULTS AND DISCUSSION**

#### DATA AND DATA GENERATION

Although I was only minimally involved in experiments, there are two reasons to describe how they were conducted. Firstly I was working on experimental design, and secondly I share the view that metadata on experiments is very important. [1]

#### DATA GENERATION

A main concept of the experimental group was to elucidate the actual *in vivo* situation and restrain using artificial factors in the setup as much as possible. This approach rises problems (see: *no clear input*), but gives some certainty, that observed response is biologically relevant (e.g pheromone is in a biologically relevant range) and minimizes the sources of possible artifacts. There is a good reason for such an approach. Vastly modified yeast species are widely used in the field, but their misbehavior is often only later observed.

- 1. Ras1 overexpression studies yielded an artifact of dual localization, giving rise to interesting, but questionable hypotheses of compartmentalized Ras1 signaling. [63]
- 2. Pombe mating studies often used an engineered strain, where cells constantly sense nitrogen starvation after heat-shock, therefore synchronicity in these cultures was great. However later was observed that mating is not correctly executed in this strain [64–66]
- 3. Many assays initiate mating by addition of synthetic pheromone. However pheromone is a very local story. It is secreted in small quantities and probably builds up a gradient around the cell, which gradient can affect behavior. A review: [67]
- 4. Also, pheromone is degraded by a secreted protease Sxa2. [68] Since synthetic pheromone is given in big quantities to reach a sufficient global concentration, the effect of Sxa2 is unknown.

#### Genetic engineering

A simple way to perturb the system is to create knockout mutant strains or to replace proteins with modified ones. All strains, inclusive the 'wild type' strain contain a GFPtag on the C-terminal of Spk1. It proved not to have observable effects on pheromone signaling, conjugation, mating or meiosis. This GFP-tag is used to monitor sub-cellular localization and as protein tag to establish absolute protein concentrations by comparing signal to a GFP-calibration curve. See the list of planned and tested mutant strains in: **Designed experiments**.

#### Protocol in the original experimental setup

The basic information source in this project was time series measurement of relative or absolute Spk1 level and relative Spk1 activity in different mutant strains. In the experiments, the assayed pombe strain is liquid cultured in full media on room temperature under gentle shaking for aeration. The culture is shifted to nitrogen-depleted media, which stops mitotic growth and prepares them for conjugation. Upon starvation cells arrest cell cycle and start to secrete pheromone, what induces pheromone response in partner cells. For detailed cellular behavior see the *Biology* chapter. In liquid culture, cells often form big aggregates, by secreting agglutinin and sticking to each other. Efficient mating happens within these aggregates.

Cells are sampled from the starvation induced culture at scheduled time points. Sampled cells are fixed; extracts are pre-purified, loaded on acrylamide gel, separated by mobility, blotted to membrane. Membranes are developed by the following fluorescent antibodies: Primary for phospho-Spk1 : *anti phospho ERK antibody : Cell signalling Phospho-p44/42 MAPK XP Rabbit monoclonal Antibody (4370S); Primary for the fission yeast alpha-tubulin: mouse monoclonal TAT1 generated by Prof Keith Gull's lab; 2ndary antibody for ppSpk1: Li-Cor IRDye 800CW Goat anti-Rabbit IgG (926-32211); 2ndary antibody for TAT1 : Li-Cor IRDye 680LT Goat anti-Mouse IgG (926-68020). Using fluorescent antibodies has a big advantage over horseradish-peroxidase-based techniques: the signal intensity is linear with the protein amount in a range of six magnitudes. On the other hand the tradeoff is that signal-to-noise ratio much lower. According to the requirements of pathway reconstruction and modeling, two kind of western blots were performed, these are described below. Western blot experiments and quantitation was done by our collaborators in Leicester.* 

#### (Semi-)quantitative western blots

For quantitative modeling, measurements were performed in multiple replicates and with high temporal resolution. Early westerns explored the approximate dynamics. Based on this, an irregular sampling was designed. The aim was to have high resolution on the dynamically rich and low resolution on dull sections. The highest resolution was 60 min, and the duration was 12 or 24 hours. Quantitation was done by the Odyssey fluorescence scanner. ppSpk- and tubulin-signal of different wawelength were background subtracted, then ppSpk1 was normalized by tubulin. Method is referred as

semi-quantitative as it showed delicate dynamical features, like the pre-activation peak, however it provided average values over a non-homogenous population.

#### Western blots for qualitative predictions

Since the project turned into a pathway reconstruction problem, the *structural model* of the pathway was used for qualitative predictions. Testing mutations did not need elaborate quantitation nor dense sampling. These westerns showed Spk1 activity of different mutant strains with 6 time points (at 0, 4, 8, 12, 24 & 36h) and were not subjected to proper quantitation. Results are provided in section: *Time course data and cell morphology*.

#### Absolute protein quantification

To deliver absolute protein concentration for the model, we applied the method used for cerevisiae before. [69] The GFP intensity of tagged proteins was measured and compared to a series of GFP standards. Thereby absolute protein count was gained for the sample. The number of cells was counted by *CASY cell counter* or by cell counter chamber under microscope. An estimate on average volume of different subcellular compartments of pombe was found in BIONUMBERS (ID: 102278). The estimates were controlled against known values for other signaling proteins and found to fall within this range. See in data supplement and [70]

#### Problems with the setup

The liquid culture setup was a robust protocol; however it yielded rather low quality data. The following problems were characterized.

1. Low mating efficiency:

Cells need a pheromone gradient (and a partner in range) for successful conjugation. The constantly shaken liquid washes away cells and gradients. Cells which cannot attach to a ball of cells, cannot mate.

- Cell synchronicity: Although starvation and pheromone distribution during shaking should synchronize cells to some extent, the starting culture is completely asynchronous.
- 3. Culture heterogeneity: There are two kinds of cells in the population. The cells in one of the 'balls' execute mating, whereas the majority is floating around alone and cannot mate. Cell extracts however cannot distinguish these cells; the westerns therefore reflect an average of two separate behaviors.
- No measurement of system input: Pheromone is not added externally at a single time point, the protocol let the cells stimulate themselves – of course this is also a source of variation.

#### Selection of synchronous cells by elutriation and plate based assay

To acquire *more quantitative* data, we tried to *elutriate* the culture. The elutriator is a special centrifuge where the cells flow in a tube from the periphery of a rotating plate towards the centre. Along the way cells separate by density, therefore it is possible to select freshly divided cells from the rest of the population. It selects a synchronous (sub) culture without externally treating the cells except the centrifugation. However, this method yet did not work out for technical reasons. My collaborators successfully developed a plate-culture based assay that allowed a high mating efficiency without the formation of aggregates, thereby heterogeneity was much smaller.

#### TIME COURSE DATA AND CELL MORPHOLOGY

The Western Blot experiments shown here are yet unpublished work of Emma Kelsall and Dr. Kayoko Tanaka at the University of Leicester, any further use or publication of the data is the exclusive right of the mentioned authors.

The Spk1 activity is the measured output of the transcriptional response. Spk1 (MAPK) is the key regulator of Ste11 transcription factor in pheromone response. Relative Spk1 activity plots and images of cell morphology of mating are presented in this section. All

Spk1 activity values are normalized by tubulin alpha intensity. The presented light microscopy pictures only show representative phenotype of the cells.

Besides measuring Spk1 levels and activity, pictures of cells were taken under light microscope. These pictures illustrated typical morphological features of mating in different mutant strains. As the model was extended to incorporate the morphology pathway, imaging validated predicted morphologies, along with westerns that tested predicted Spk1 activity. Three, focal strains represent the most important morphological and genetic behavior types. These are the wild type, the hypersensitive *Ras.Val* and the downregulation deficient *Byr1.DD*. These three strains were quantitatively modeled. Data on other mutants was used for pathway reconstruction.

For experimental details, see: *Data generation*. Quantified original data, comparisons and western blot quality control stands in the supplementary excel file.

#### Wild type strain

Wild type behavior is written in detail in the *Biology* section. The wild type Spk1 activity shows the following features. It climbs to reach a transient peak, then it falls back to a non-zero level, from there it gradually decreases slowly. With the current protocol it is very hard to estimate how long this gradual decrease takes since the culture is a mix of mating and not-mating cells. See: *Problems with the setup*.



FIGURE 4: SPK1 ACTIVATION PROFILE OF WILD TYPE CELLS IN TWO BIOLOGICAL REPLICATES. THE AVERAGE PEAKING TIME IS AROUND 7 HOURS. © EMMA KELLSALL & DR. KAYOKO TANAKA

After comparing many independent experiments it became clear that the early 'shoulder' of the main activation peak around 2-3 hours is not an artifact. It also appeared in the *Ras.Val* and the *Byr1.DD* strains. The hypothesis, the competing models and designed experiments are to be seen in: *Models for subsequent activation peaks*.



FIGURE 5: VEGETATIVE AND MATING WT CELLS. MEIOSPORES ARE ALREADY FORMED IN THE CELLS AS A RESULT OF SUCCESSFUL MATING. WILD TYPE CELLS NEVER FORM LONG SHMOOS AS *RAS.VAL* CELLS DO. © EMMA KELLSALL & DR. KAYOKO TANAKA

Wild type cells in vegetative cell cycle show an elongated tubular form. Newly divided cells often remain together.

#### Ras1.G17V strain

The *Ras1.G17V* strain contains a mutation in the regulatory domain of Ras1 that inhibits the binding of Gap1<sup>7</sup>, a GTP->GDP exchange factor. In further the strain is referred as 'Ras.Val'. Mechanically, once Ras1 is activated (GTP binding) it is only deactivated by its own, comparably slow GTPase activity. The mutation is confirmed by the fact that cells lacking Gap1 behave the same way.

The strain is called 'pheromone hypersensitive': the cells grow thicker and longer shmoos after stimulus. This strain also shows a higher and earlier phospho-Spk1 peak. The strain is sterile; shmoos do not reach partner cells. Whether they are incapable of fusion, or they cannot find the direction of the partner, or just mistake in estimating the partners' direction is to be enlighten by further experiments. After looking at microscopy images, I set up a hypothesis that claims that cells can sense the direction, but start shmooing towards any distant partner. This concept is elaborated in *The distance measurement hypothesis* section.

One can observe 3 features of Spk1 activity comparing to wild type response: 1.These cells fires stronger. 2. These cells fires earlier. 3. The signal is downregulated earlier. The first feature can be plausibly explained by the molecular mechanism; why the pathway fires earlier is not so plausible. *Ras.Val* strain has higher background activity compared to WT (see: *Unquantified western blots*). It is possible, that the main peak is a result of autocatalytic self activation, since at least six positive feedbacks (Mam2-, Gpa1-, Ste6-, Ste4-, Spk1- and Ste11-transcription) exist in the pathway. Systems with coupled positive and negative feedback loops can be bistable, and the pombe pheromone response is such a system. Once they cross a critical activation threshold, the autocatalytic effect grows stronger and stronger and the system falls into the active state. Now since *Ras.Val* cells have higher Ras1 basal activity, they are closer to this threshold. This hypothesis is also supported by the final full scale model.

The picture in the downregulation is somewhat clearer. The hypersensitive *Ras molecule itself* is slower downregulated than WT, since it lacks interaction with Gap1. Strikingly, the strains activity falls earlier. This contradiction could be resolved by the Spk1 triggered negative feedback. Since it is earlier and stronger upregulated, it earlier

<sup>&</sup>lt;sup>7</sup> Gap abbreviation stands for G-protein Activating Protein.

and stronger activates the downregulating mechanism. The faster Spk1 downregulation with a slow deactivated Ras1 molecule may also suggest that the main downregulation is downstream of Ras1. This conclusion makes less likely, that the typical regulation by receptor internalization and/or desensitization is the key regulator here. [71] Since this strain is sterile and still cuts the signaling off, the downregulation cannot depend on successful conjugation, it must be activated by the pheromone pathway, regardless of the outcome of mating.



FIGURE 6: THREE BIOLOGICAL REPLICATES OF THE RAS.VAL STRAINS SPK1 ACTIVITY. THE AVERAGE PEAKING TIME IS AROUND 4.67 HOURS; THIS FASTER RESPONSE IS CALLED THE EARLY SHIFT. THIS IS PARTLY SEPARATE FROM THE OTHER EFFECT, THE HIGHER INTENSITY OF THE RESPONSE. WHEREAS THE FIRST IS THOUGHT TO BE THE DIRECT RESULT OF THE HIGHER BASAL ACTIVITY, THE SECOND IS THE RESULT OF THE HIGHER SATURATING ACTIVITY OF RAS1 IN THE RAS.VAL STRAIN. © EMMA KELLSALL & DR. KAYOKO TANAKA



FIGURE 7: VEGETATIVE RAS.VAL CELLS LOOK AS THE SAME AS WILD TYPE CELLS, SUGGESTING THAT RAS1 HYPERSENSITIVITY DOES NOT AFFECT VEGETATIVE GROWTH SERIOUSLY. RAS.VAL CELLS RESPONDING TO PHEROMONE LOOK DIFFERENT THAN WT CELLS, THEY GROW ELONGATED AND THICKER SHMOOS; HOWEVER THEY ARE INCAPABLE OF SUCCESSFUL MATING. © EMMA KELLSALL & DR. KAYOKO TANAKA

#### Byr1.DD strain

*Byr1.DD* is a phosphomimetic protein mutant, which constitutively active, and cannot eb further activated by phosphorylation. Byr1 has two sites phosphorylated by Byr2, which are both replaced by Aspartic acid [D] in the mutant protein. The negative charge mimics the effect of the similarly negative phosphate-group. The major feature of pheromone response in *Byr1.DD* cells is that the signal is never switched off. This means that the downregulation must happen upstream of Spk1, but where exactly is unknown. Morphologically, cells grow slightly bigger shmoos than wild type; they localize and attach to their partners correctly, but subsequent fusion is missing. Comparing this with the apparent loss of downregulation we hypothesized that shutting down the transcriptional response is probably necessary for the expression or de-regulation of a protein involved in fusion. To support this idea, collaborators are generating a strain containing both *Byr1.DD* mutation and analogue-sensitive Spk1. In this way, Spk1 can be regulated externally. The experiment is still to be done. Our expectation is that the artificial downregulation of the signal restores the WT phenotype.



FIGURE 8: SPK1 ACTIVITY OF BYR1.DD STRAIN IN TWO BIOLOGICAL REPLICATES. THE SATURATION TIME IS DEFINITELY NOT BEFORE 12 HOURS, AND IT REACHES CONSIDERABLY SMALLER VALUES THAN WT DOES. THE EARLY SHOULDER OF THE ACTIVATION PEAK APPEARS AS IN MOST OTHER MUTANTS. © EMMA KELLSALL & DR. KAYOKO TANAKA

Once *Byr1.DD* protein is expressed, it activates the pathway: since there is no signaling before starvation, either Byr1, Spk1 or both are not expressed upon nitrogen depletion. The regulation of Byr1 is unknown, possibly due to its very low molecule count.

Therefore we cannot exclude any of the possibilities. Spk1 is actively induced by starvation. [31] Our experiments showed that it is below measurement accuracy before starvation-induced stimulation.

Since Byr1 cannot be phosphorylated by Byr2, it also means, that the downstream pathway does not sense any input, so this strain is insensitive to upstream perturbations. Input-insensitivity was also confirmed by *Byr1.DD* 'starvation-only' experiments, which showed the same phospho-Spk1 profile as cells induced by *starvation and pheromone*. The cells nevertheless did not grow shmoos, since the morphology pathway was not stimulated.

*Byr1.DD* mutation also proved to be a powerful tool for highlighting the function of pathway components: combinational mutants containing a knockout in upstream components still had a functional (up to before fusion) MAPK-cascade. In this way we could show if a component is important for the morphology pathway. If the knockout in *wild type* showed a reduced Spk1 activity, the examined protein is also an effector the transcriptional response.



FIGURE 9: MATING BYR1.DD CELLS ARE SIMILAR TO WILD TYPE: THEY FORM SMALL SHMOOS, AND CORRECTLY ATTACH TO A PARTNER; HOWEVER THEY CANNOT FUSE WITH THE PARTNER. © EMMA KELLSALL & DR. KAYOKO TANAKA


#### $sxa2\Delta$ strain

FIGURE 10: SPK1 ACTIVITY IN THE SXA2Δ CELLS (TWO BIOLOGICAL REPLICATES). THE PEAK TIME IS AROUND 12 HOURS. ALTHOUGH THE SIGNAL IS DOWNREGULATED (OPPOSED TO BYR1.DD), IT HAPPENS MUCH LATER THAN IN OTHER STRAINS. © EMMA KELLSALL & DR. KAYOKO TANAKA



#### gap1∆ strain

FIGURE 11: SPK1 ACTIVITY IN THE GAP1Δ STRAINS. THE PEAK TIME IS AROUND 4 HOURS. THE GAP1Δ STRAIN IS THOUGHT TO HAVE THE SAME FUNCTIONAL LOSS AS RAS.VAL MUTANTS: THE INTERACTION OF RAS1 AND GAP1 IS MISSING. © EMMA KELLSALL & DR. KAYOKO TANAKA



FIGURE 12: MATING PHENOTYPE OF GAP1 $\Delta$  CELLS. THE CELLS GROW LONG LINEAR SHMOOS AS RAS.VAL CELLS DO, AND THEY ARE INCAPABLE OF MATING TOO. THE CELLS THOUGH LOOK A BIT MORE 'DOTTY'. © EMMA KELLSALL & DR. KAYOKO TANAKA

#### Remarks

Error bars are not reported for the following reasons: in many cases there are just two biological replicates done, while these are normally done in three technical replicates, we are interested in biological variance, which is not represented in technical replicates, therefore it could be misleading. Addressing biological variance is not trivial, since variance in peaking time and amplitude could be caused by different reasons. Treating the two variances together is misleading.

For example the two *Ras.Val* cultures below show little variance at 5 hours in Figure 5, however it does not mean that they have little variance in peak amplitude. One should consider treating the time- and amplitude- variance of the data separately. One way would be to establish the variance in peaking time, than shift the datasets so that they have the same imaginary peaking time. Finally one could establish the variance in peak amplitude, which is now not biased by variance in by the response offset of different cultures.

# **Unquantified western blots**

Unquantified or *-quick and dirty-* western blots provided qualitative information for the reconstruction of the signaling pathway.



FIGURE 13: ORIGINAL IMAGES OF WESTERN BLOTS. THESE UNQUANTIFIED WESTERNS WERE USEDTO TEST QUALITATIVE PREDICTIONS. GREEN INDICATES THE FLUORESCENCE OF ANTI-PHOSPHO-SPK1 ANTIBODY, RED THE ANTI-TUBULIN ANTIBODY. NUMBERS OVER COLUMNS DENOTE THE HOURS AFTER STARVTION. 1<sup>ST</sup> PANEL: OVERLAY IMAGE OF BOTTOM PANELS. 2ND PANEL: THE UPPER BARS SHOW PPSPK1 (ANTI-PPERK), THE LOWER IS AN UNSPECIFIC OFF-TARGET SIGNAL OF THE ANTIBODY. THE TWO ADJACENT ROWS OF TUBULIN (ANTI-TAT1) REPRESENT TWO ISOFORMS. 3<sup>RD</sup> PANEL: THE UPMOST BAR MARKS THE TOTAL SPK1 (ANTI GFP), WHEREAS THE LOWER ADJACENT ROWS MARK TUBULIN (ANTI-TAT1). © EMMA KELLSALL & DR. KAYOKO TANAKA

39 / 120

Experiments from September and October 2012



FIGURE 14: ORIGINAL IMAGES OF WESTERN BLOTS DONE IN SEPTEMBER. THESE UNQUANTIFIED WESTERNS WERE USEDTO TEST QUALITATIVE PREDICTIONS. GREEN INDICATES THE FLUORESCENCE OF ANTI-PHOSPHO-SPK1 ANTIBODY, RED THE ANTI-TUBULIN ANTIBODY. NUMBERS OVER COLUMNS DENOTE THE HOURS AFTER STARVTION. TOP PANEL: THE UPPER GREEN BARS SHOW PPSPK1 (ANTI-PPERK), THE LOWER IS AN UNSPECIFIC OFF-TARGET SIGNAL OF THE ANTIBODY. THE TWO ADJACENT ROWS OF TUBULIN (ANTI-TAT1) REPRESENT TWO ISOFORMS. BOTTOM PANEL: THE UPMOST RED BAR MARKS THE TOTAL SPK1 (ANTI GFP), WHEREAS THE LOWER ADJACENT ROWS MARK TUBULIN (ANTI-TAT1). © EMMA KELLSALL & DR. KAYOKO TANAKA:

The absence of signal in  $gpa1\Delta$  strain is striking. (Only the lower green bar is apparent, that is possibly a ppSAPK, an off-target antibody signal) The signal is much lower than in case of any other knockouts. I could conclude two possibilities from that, which are not mutually exclusive. Either Gpa1 activation is required very downstream; therefore its loss filters out basal pathway activation. Alternatively it is possible that Gpa1 is *obligatory for starvation induction*. The first idea is implemented in the final model. Our concepts about Gpa1 are yet to be clarified. A *Byr1.DD & gpa1A* strains 'starvation only' experiment could provide further insight, but may not be decisive. A few years ago, it was proposed that Ste6 and Efc25 act strictly complementary to each other. [72] We could not reproduce these findings. Our result shows that the loss of Ste6 does not abolish signaling in the transcriptional branch; therefore it cannot be the only activator of that pathway. For this reason, Efc25 and Ste6 may act in o the same way: charge Ras1 with GTP.

# *GENOME WIDE PROTEIN EXPRESSION DATA* **Dataset**

We were kindly provided with a genome-wide dataset by Jürg Bähler of protein and mRNA concentrations on 2 conditions. [62] -in press- Cells from nitrogen-rich media (vegetative cell cycle, abbreviated MM) and quiescent cells from 24h N-starved culture were measured (abbreviated MN). The quiescent dataset was measured in cells 24 hours after N-starvation, when cells go to a dormant state, instead of mating followed meiospore formation in our experimental setup. Our experiments and the model start from vegetative stage, so the first condition provides usable data. The dataset contained information for special, non-coding or unknown DNA sequences, which were removed: antisense RNA, "dubious" sequences, 5s rRNA, tRNA, small nucleolar RNA, sequence orphan, pseudogene, non-coding RNA (predicted), 18S ribosomal RNA, intergenic rna, 5.8S ribosomal RNA, 28S ribosomal RNA, 8S rRNA (predicted), 7SL signal recognition particle component, LTR - assumed long terminal repeats, SPBC1348.05, SPBC28E12.02, ght7, SPBPB8B6.02c, mug27, mde7, SPBPB2B2.07c, pma2, mug28, mug62, , RNase P K-RNA, telomerase RNA, translationally silent transcript from tco1 locus, RNAse MRP, 21S rRNA, C/D containing snoRNA Z30, 12 small nucleolar RNA/small nucleolar RNA snR99, 20 small nucleolar RNA, SnoRNA (predicted), Tf2-7, Tf2-fragment5. After filtering, 4747 measurements remained.

# Estimator

The model could make use of 10 protein's levels (cpc, average count per cell) from the dataset: byr1, rgs1, sxa2, ste4, ras1, spk1, mam2, byr2, gpa1, ste6. Some of them were directly measured, others protein counts were estimated from other provided data: *mRNA level, Protein isoelectric point [pH], Protein mass [kDa], Ribosome occupancy, Ribosome number, Ribosome density, Relative poly(A) length, Relative PolII occupancy, mRNA half-life.* All of these quantified properties and the single products with *mRNA counts* were tested for correlation with protein levels. All showed a worse correlation than mRNA level itself. Strikingly, mRNA half-life showed no correlation with protein levels. The same result was found in: [73] Satisfyingly, mRNA counts showed an acceptable linear correlation and yielded the following equations:

 $mRNA\_count (MM) * 1449 = protein\_count(Vegetative)$ (1.1.) The linear equation had a coefficient of determination (R<sup>2</sup>) of 0.619.

 $mRNA\_count (MN) * 3696 = protein\_count(Quiescent)$  (1.2.)

The linear equation had a coefficient of determination ( $R^2$ ) of 0.49. This dataset was not used in the model.

The correlation was tested for linearity, and generally worse higher-order or sub-linear correlation were found. Although a slightly better linear fit ( $R^{2} = 0.623$ ) was achieved with the correlation between *mRNA-count*<sup> $\Lambda_{1.1}$ </sup> and protein count; however the difference was too slight to change the simple linear correlation model.

# **Quality of estimate**

After establishing a linear correlation, it was tested for quality in predicting protein concentrations in both conditions (vegetative growth, quiescent). Correspondingly 2443 and 2646 protein measurements were available in the dataset, where the prediction could be compared. In 98% and 96% of the cases the estimate was within one magnitude, and in 37.0% and 37.3% of the cases, the estimate was better than +/- 50%. This was very satisfying result, knowing that often mRNA and protein levels often do not correlate at all. Nevertheless it was apparent that Ste11 dependent proteins are very different to other proteins. Most of them are transcribed in very small quantities and no translated protein was found (NA).

Common n		MN mRNA	MM protein	MN protein	Estimated
ame	cnc	cnc	cnc	cnc	mM protein
ame	.cpc	.cpc	.cpc	.cpc	срс
sxa1	12.00	4.00	3673	9031	17388
cdc42	10.00	2.60	29312	21374	14490
shk1	4.00	1.10	4164.45	1627.98	5796
scd2	3.20	1.30	901	2081	4637
efc25	3.20	1.10	99	NA	4637
byr1	2.70	0.84	3354	NA	3912
ras1	2.20	0.44	9899	7003	3188
scd1	2.10	0.81	593	454	3043
ste11	1.80	1.70	NA	NA	2608
gpa2	1.60	0.39	1995	2086	2318
byr2	0.88	0.30	NA	NA	1275
gpa1	0.72	0.67	1981	2381	1043
sxa2	0.66	0.19	NA	NA	956
spk1	0.59	0.89	NA	NA	855
ste4	0.57	0.26	343	NA	826
rgs1	0.38	0.14	NA	NA	551
ste6	0.25	0.21	NA	NA	362
map2	0.16	0.02	NA	NA	232
mam2	0.06	0.16	NA	NA	84

TABLE 1: A SUBSET OF GENES FROM THE GENOME-WIDE DATASET [62]. MM DENOTES VEGETATIVE GROWTH, MN QUIESCENT STATE (24H AFTER N-STARVATION). THE 1<sup>ST</sup> AND 2<sup>ND</sup> COLUMN SHOWS AVERAGE MRNA COUNTS PER CELL (CPC), 3<sup>RD</sup> AND 4<sup>TH</sup> COLUMN PRESENTS THE MEASURED PROTEIN COUNT PER CELL. THE 5<sup>TH</sup> COLUMN CONTAINS THE ESTIMATED PROTEIN COUNT CALCULATED AS DESCRIBED ABOVE. CLUSTERING OF PROTEINS RELATED TO PHEROMONE RESPONSE IS DENOTED BY BACKGROUND COLOR. PROTEINS COLORED IN <u>GREY</u> ARE CURRENTLY NOT, OR NOT SEPARATELY PART OF THE MODEL, <u>GREEN</u> MARKS THE STE11 DEPENDENT (TRANSCRIPTIONALLY REGULATED) GENES, WHEREAS <u>WHITE</u> COLORED LINES CONTAIN GENES THAT ARE TRANSCRIPTIONALLY <u>NOT REGULATED</u>. IT IS APPARENT THAT MOST STE11 DEPENDENT PROTEINS HAVE VERY LOW MRNA COUNTS, WITH NO PROTEIN SIGNAL EXCEPT GPA1 AND STE4. NOTE THAT SXA1 AND SXA2 HAVE VERY DIFFERENT FUNCTIONS DESPITE THEIR SIMILAR NAMING<sup>8</sup>.

# Application of data

Measured protein counts were used for the model as pre-stimulation steady state data in 4 cases (ras1, byr1, ste4, gpa1), and estimates were utilized in two cases (byr2, ste11). Total Spk1 counts were estimated from this dataset but also measured by our collaborators. This gave a chance of validation. The estimate was reasonably close to the measured value. 855 counts per cell (cpc) were estimated, while 600 and 400 cpc was measured by the collaborators. The average of the measured values was used for the model (500 cpc).

The absolute protein counts were transformed into concentrations (uMol/liter) with a short R-script, which makes use of a microscopy based cytoplasm- and nucleus- volume estimates. [70] Proteins were distinguished by their localization for the calculation of

concentration: cytoplasmic, or cytoplasmic and nuclear. (No proteins resided in the nucleus only). Concentration for extracellular protein Sxa2 was not used. Localization was established by the online 'S. pombe Postgenome Database' database [10] and confirmed by literature. For the current state of the model no separate compartment is treated for proteins localized at the shmooing location or at the membrane in general. The calculated protein concentrations were in the range of other measured signaling proteins in pombe (the cytokinesis proteins: 0.09-4.75u Mol/l). [70] The only exception is the pheromone receptor, which is estimated (and assumed in literature) to be almost absent in vegetative state.

# Improvements of estimate

# Manual curation of the dataset

The estimate of protein counts could be improved in two ways. The more trivial way is to filter the dataset further. Signaling proteins appeared to have characteristically different translational behavior than other proteins. It was also found that mRNAs and proteins involved in fast responses have lower stability (half-lives) in mammals. [73] Initial filtering of the dataset increased the quality of the linear regression particularly; therefore further filtering may also be useful.

# Probability distribution based models for gene expression

A second way to increase prediction quality for protein levels is to make use of the originally measured distribution of protein and mRNA counts, instead of using mean values. Based on these data, a probability function can provide mapping between the distribution of mRNA counts and the distribution of protein counts. [74,75] These are on the other hand quite complicated methods using way more sophisticated models for translation than linear. Due to the limited time, these improvements were not applied.

# Cross validation

A further notice is that the quality control of estimates is not statistically correct. A correct way would be to apply cross validation by splitting up the data into training and validation subsets. One suitable method, termed *10-fold cross validation* could be applied for the problem. In this, the dataset is split up into ten subsets. In an iterative manner, each time different 9 subsets are selected for training and the remaining one for validation. The average performance of predictive power is reported.

# PATHWAY RECONSTRUCTION RESULTS

## Starting point

The project started with a rather simple pathway scheme shown in Figure 2. This scheme represents the initial notion of the pathway. The depicted interactions are unspecific: arrows and associations mean various functions. It was not considered to view the system as a genetic regulatory network. The considered network is very simplistic, but from a *dynamical systems* view, the considered species are not always the most important ones: pmp1, pyp1, gap1. [41,87] A number of components were ignored (Ste4, Ste6, Shk1 and the Cdc42 complex), some of them turned out to be important. The role of nitrogen starvation was unclear. Similarly, some other basic concepts were not explicitly formed. My contribution was partly the explicit (often mathematical) formulation of the design motifs, or hypotheses about them.

#### **GENETIC REGULATION**

The original figure (Figure 2) depicts the pathway as a usual post transcriptionally regulated signaling network. In such a pathway the dynamics can be explained by subsequent post-transcriptional activation and deactivation steps. Although PTM's are ubiquitous in the pathway, the timescale of the response (6-12h) suggested that the regulation happens on the level of gene expression. It indeed turned out to be the case. Nine genes are significantly induced (map2, mam2, sxa2, gpa1, rgs1, ste4, ste6, spk1, ste11) whereas six remain transcriptionally unregulated (ras1, byr2, gap1, pmp1, pyp1<sup>9</sup> and possibly byr1,). See in *Transcriptional regulation*; inferred from the Bähler Expression viewer.

# STARVATIONAL INDUCTION

It was known that starvation prepares the cell for pheromone signaling by transcription of some of the components. [12] It is also known that this is mediated by Ste11. However it was not known whether starvation signal actually converges to the pheromone pathway and if it causes Spk1 activation.

Looking more carefully at the activation profile (*Time course data and cell morphology*, data supplement) it is apparent that activation happens in two rounds in the majority of the measurements. In most of the data, a small initial and a big

<sup>&</sup>lt;sup>9</sup> Pyp2 is a known phosphatase regulated by Ste11 activity, also a homolog of pyp1; but it was shown not to be important to the pheromone pathway. [87]

subsequent peak are visible. See Figure 15 as example in case of the *Byr1.DD* strain. Two alternative hypotheses about the pre-activation are detailed in section: *Models for subsequent activation peaks*. From network reconstruction perspective it is important where the input reaches the pheromone pathway.



FIGURE 15: ZOOM-IN TO THE INITIAL STEPS OF BYR1.DD STRAINS SPK1 ACTIVATION PROFILE. ON HIGHER TEMPORAL RESOLUTION WE MAY FIND OUT MORE ABOUT THE EARLY SIDE PEAK.

I proposed an experiment where WT cells sense starvation only. The measurements have shown a minor early activation, supporting the hypothesis that starvation causes Spk1 activation, not only Ste11 mediated transcription. (Figure 12, panel A) This means that the starvation signal utilizes at least partially the pheromone pathway. I proposed another experiment, *Byr1.DD* - starvation only. Spk1 behaved as in Byr.DD cells with pheromone (Figure 12, panel C). The experiment supported the idea that starvation activates Spk1, although we did not know how exactly. Since *Byr1.DD* mutation cuts off the signal pathway so that everything upstream of *Byr1.DD* is indifferent. It means that the starvation-mediated activation is localized downstream of Byr1, namely it acts at least either on Spk1 or Ste11 or both. The Byr1.DD starvation only experiment had not enough temporal resolution to show whether Spk1 shows a dual activation peak upon a single starvational stimulus<sup>10</sup>. If this is the case, it suggests that both self sensitization

<sup>&</sup>lt;sup>10</sup> Nevertheless Byr1.DD cells always sense pheromone by the constitutively active Byr1.

and dual activation can play a role in the observed dual peaking. A decisive experiment could be done by using a non-functional Spk1 protein.

Later questions rose whether the experimental setup guaranteed that cells do not sense pheromone, therefore the experiments are to be modified and repeated. (By receptor deleted strains). For this reason these results are not incorporated in the final model.

## Full picture and re-simplification with a different focus

Instead of modeling with the initial scheme, I aimed to give a full description, that was later re-simplified (as much as possible) with a different focus at the end. First, I started to assemble the full interaction network from interaction databases (Smart, BioGRID), literature and the experiments.

It became clear that some of the considered proteins are not important regarding the qualitative shape of the dynamics. These were: Gap1, Pmp1, Pyp1, and maybe Ste6. Whereas the initially neglected Scd1 activated complex turned out to be necessary for signaling. Finally we achieved a picture that contained all components that qualitatively shape the dynamics, but disregard those do not. (Faded in Figure 16)

The available knowledge was however not enough to come up with a definite picture. All the functionally important proteins are likely to be known, since these would have appeared in mutagenic screens. However the relation among them is much less clear. This is also mirrored in the scheme proposed in this work. There are three cores of vagueness. One is Gpa1, whose effectors are completely unknown, and the possible other roles of Gpa1's GEF, Rgs1. The other unclear spot is Ste4, its activation and effect on Byr2. [88] Gpa1 and Ste4 are connected in the presented scheme; however this is only a hypothesis. The third unclear point is the complex regulation of the transcription factor Ste11. Luckily, most of Ste11 regulation seems not necessary for explaining our data, so this issue seems to have the smallest impact on the models prediction.



FIGURE 16: FINAL NETWORK STRUCTURE CONSISTENT WITH THE LATEST DATA (OCTOBER). THE GPA1-STE4 INTERACTION WAS NOT PROVEN, IT IS JUST ONE POSSIBLE HYPOTHESIS FOR THE OBSERVATION THAT GPA1 FUNCTIONALLY ACTIVATES SPK1 INDEPENDENTLY OF RAS1. THE LITERALLY ZERO SPK1-ACTIVITY IN GPA1Δ STRAINS SUGGESTS THAT IT ACTS CLOSE UPSTREAM OF SPK1; IN ITS ABSENCE FURTHER UPSTREAM NOISE DOES NOT REACH SPK1. FOR A HORIZONTAL, BUT SMALLER PRESENTATION OF THI FIGURE SEE: FIGURE 20.

# EXPLAINING SPK1 ACTIVITY AND MORPHOLOGY IN THE SAME MODEL

Another main advance of my work is to treat morphological and transcriptional response in the same model. I looked at cell morphology as a source of useful information that is never exploited before to understand the transcriptional response. It was proposed that Shk1, a member of the morphology pathway is an activator of the transcriptional branch. [52,89] This however was never proven *in vivo*, nor is it discussed intense in literature. Charting the regulatory loops also pointed out that the transcriptional branch indirectly regulates the morphology branch, therefore it is a vice versa relation: Figure 17.

I proposed and experiment to show that an intact morphology pathway is necessary for the activity of the transcriptional branch. However it is not straight forward as Shk1 as well and its probable activator, Cdc42 are essential genes, therefore cannot be knocked out. [90], reprinted also in "01\_Data\_supplement/03\_viability" It is known that Shk1-Scd1-Scd2-Cdc42 is a functional complex, therefore my collaborators suggested that disruption of any components may disrupt the activity of this complex. This idea worked out nicely, and they have shown that  $Scd1\Delta$  strain showed a strongly diminished Spk1 activity and no shmoo was grown.



FIGURE 17: ROUGH SCHEME OF THE PHEROMONE RESPONSE DEPICTING THE BRANCHING STRUCTURE OF THE PATHWAY, THE CROSSTALKS AND MAIN FUNCTIONS.



FIGURE 18: SCD1Δ WESTERN BLOT COMPARED TO WT RESULTS (FROM DATA SECTION). ALTHOUGH THE EARLY PEAK AROUND 4 HOURS IS PRESENT IN BOTH STRAINS, NO MAIN ACTIVATION PEAK IS APPARENT. NUMBERS OVER COLUMNS DENOTE THE HOURS AFTER STARVTION. 1ST PANEL: OVERLAY IMAGE OF BOTTOM PANELS. 2ND PANEL: THE UPPER BARS SHOW PPSPK1 (ANTI-PPERK), THE LOWER IS AN UNSPECIFIC OFF-TARGET SIGNAL OF THE ANTIBODY. THE TWO ADJACENT ROWS OF TUBULIN (ANTI-TAT1) REPRESENT TWO ISOFORMS. 3<sup>RD</sup> PANEL: THE UPMOST BAR MARKS THE TOTAL SPK1 (ANTI GFP), WHEREAS THE LOWER ADJACENT ROWS MARK TUBULIN (ANTI-TAT1). © EMMA KELLSALL & DR. KAYOKO TANAKA

We have not proven the physical interaction as before *in vitro*<sup>11</sup> (Shk1 -> Byr2 activation), instead the Shk1-Scd1-Scd2-Cdc42-> Spk1 functional activation. Still after this relation was shown, we could use simple morphology to speculate about activity in transcriptional branch and vice versa.

# AN EXPLANATION FOR RAS1

The initial focus of the experimental group was Ras1, its function and its effect of its mutation on transcription. Although my collaborators showed that Ras1 is necessary for mating (See: Figure 13), and it is a central node in the network, it is surprisingly found not to activate any of its substrates. [52] It is unclear what the concrete molecular function of Ras1-GTP is. It is known to bind soluble proteins Cdc42 and Byr2 at the membrane. One of its effectors is Byr2. It has a regulatory domain that binds the kinase domain, thereby inactivating it until an the kinase domain is released. One study showed that Ras1-GTP itself cannot open up Byr2, which means that the Byr2 kinase needs to be activated additionally. [52] Shk1 was shown as a potent Byr2 activator by the same study.

The only proven role of Ras1 seems to be localizing Byr2 to membrane regions where Ras1 is activated. Why the localization of Byr2 is important is not explained yet. I came up with the hypothesis that Ras1-GTP neither activates Byr2, nor simply localizes it;

<sup>&</sup>lt;sup>11</sup> Two hybrid assay in cerevisiae.

rather the colocalization of Byr2 and the Cdc42-Scd1&2-Shk1 complex is the crucial step. By their colocalization their interactions can be enhanced. In this hypothesis Ras1 basically concentrates the two components on a small area of the membrane, thus increasing their interaction by magnitudes. It still does not exclude the possibility that Ras-GTP binding of Byr2 leads to a conformation where Byr2 is more active or easier to activate, it only suggests, that a key contribution to the pathway activation can be the colocalization of these two components.

This hypothesis seems compatible with other available data, and it explains it without assigning any unknown functions to Ras1 or other proteins. A simplistic, nonparameterized model showed that it *can be sufficient* to explain a pheromone triggered activation of Spk1. See in: *The Ras1 colocalization model*. Although Ras1 is essential for functional Spk1 activity, the newest experiments showed that Gpa1 has a Ras1 independent way to activate Spk1. (1<sup>st</sup> row in: Figure 13) The newest results shown literally zero Spk1 activity in *gpa1* $\Delta$  strains, which suggests that Ras1 mediated colocalization of Shk1 and Byr2 is insufficient for pathway activation. Nevertheless Ras1 was also shown to be necessary for above-basal activation of Spk1. It is known that Byr2 has more than one steps of activation: conformational change and phosphorylation; both seems necessary. [52] Ste4 mediated dimerization has an unknown role. A reasonable hypothesis incorporates the new data is that Ras1 activates Byr2 by colocalization with Shk1, but Gpa1 mediated downstream activation of Byr2 is also necessary. (Figure 20)

# **DESIGNED EXPERIMENTS**

As the part of the pathway reconstruction, I designed more than 17 experiments of which 11 have been carried out so far. The brief summary of proposed experiments, including prediction and validation (in blue) and falsification (in red) is presented in Table 2 and Table 3. A detailed list of experiments can be found in the supplementary part.

#	Strain	Condition and	Predicted Spk1	Predicted	Explanation / Conclusion
		Culture	activity	morphology	
Exp	Experiments performed beforehand		d Observed Spk1 activity	Observed shmoo morph	ology
	wt	Starvation + h90	Transient peak before	Short shmoos, correct	Conclusion: The transien peak means there
		Pheromone	downregulation. Early	attachment, than	isa time window for Spk1 activity.
			sidepeak.	fusion, than	
				melospore formation	
	byr1.DD	Starvation + h90	Delayed saturation, no	Short shmoos, correct	Conclusion: The downregulatory
		Pheromone	downregulation. Early	attachment, but <b>no</b>	mechanism does not act directly on Spk1.
	ras val	Stanuation + b00	sidepeak.	fusion	It can be anywhere upstream.
	ras.var	Pheromone Not	transient neak before	linear shmoos no	interaction RasVal has a higher hasal
		riteronione	earlier	attachment	activity and a higher maximal activity:
			downregulation. Early		these cause earlier and stronger peak, with
			sidepeak.		earlier downregulation.Since these cells do
					not mate, the downregulation is not a
					result of sucessful mating.
	gap∆	Starvation + h90	Early & strong	Long and thick always	Conclusion: Cells do not have Ras1-Gap1
		Pheromone	transient peak before	linear shmoos, <b>no</b>	interaction. The same results as above
			earlier	attachment	confirm our knowledge about Ras1-Gap1
		Chammelian I hou	downregulation. Early	Long and thick always	interaction.
	SX020	Pheromone	transient neak before	linear shmoos <b>no</b>	downregulation of the signal hower it may
		riteronione	much delaved	attachment	be involved in early downregulation.
			downregulation. Early		
_		_			
1st	round of ne	ewly designed exper	m Predicted Spk1 activity	Predicted shmoo morph	nology & observation
01	wt	Starvation h-	No main peak, but	No shmoos	Explanation: Nothing stimulates the
			starvation induced		morphology pathway or the threshold for
			No main peak, but small	e No shmoos	Conclusion: Starvation does not result in
			···· / ··· / · ··· / · ··· /		shmoos.
02	ras.val	Starvation h-	No main peak, but	No shmoos	Explanation: Ras.Val basal stimulation of
			starvation induced		the morphology pathway is not enough to
			early "side"-peak		commit to shmoo.
			No main peak. but small	e No shmoos	Conclusion: RasVal basal activity is not
		_			enough for shmooing.
03	byr1.DD	Starvation h-	Same as h90: delayed	No shmoos	Explanation: The activity of the GeB and
			saturation, no		subsequent indirect GeB -> MoB crosstalk
			downregulation		by transcription is not enough to induce
			Same as h90: delayed sa	ti Na shmaas	shmooing. Conclusion: Starvation does not result in
			Sume as not. actayed su		shmoos, even if the transcriptional appr.
		_			program is correcty working.
04	scd1∆	Starvation + h90	No / very low (noise	No shmoos	Explanation: Activation of Byr2 by Pak is
		Pheromone	mediated) activity		necessary for pathway function. In the
					scd1? strain the Scd1&2-Cdc42-Pak1
					complex cannot assamble, therefore Cdc42
		-	No / very low activity. Ea	r. No shmoos	Conclusion: Scd1 is necessary for GeB

Proposed experiments

TABLE 2: DESIGNED EXPERIMENTS 1. FULL TABLE IS AVAILABLE IN THE SUPPLEMENT.

activity at least; shmooing anyway never appears if the transcriptional program does

not work.

#	Strain	Condition and		Predicted Spk1	Predicted	Explanation / Conclusion	
200	round of n	Culture		activity Predicted Spk1 activity f	morphology	checonyotion	
200		ewiy designed e	sperie		No shree os	Surlayetion Nething on estinate Deel in	
06	STE64	Pheromone	N90	absent.	No snmoos.	appropriate amount, therefore no Ras1 m mediated signaling happens.	
		_		Only slightly weaker activity	No shmoos	Conclusion: Ste6 does not seemunique GEF, probably Efc25 can also be utlized to stimulate the GeB. Ste6 activity is mediated by Ras1, since Ras1 negative strain is much severe	
07	ras1∆	Starvation + Pheromone	h90	No / very low (noise mediated) activity	No shmoos.	Explanation: Ras is the only way how pheromone stimulation could reach downstream targets.	
		-		Low but not zero	No shmoos,	Conclusion: There exists a Ras1	
				activity, it is definitely above noise level.	completely round cells already in vegetative state.	independent signaling to Spk1, however it is not enough to activate Spk1 to be functional. The model structure has to be	
08	gpa1∆	Starvation + Pheromone	h90	No / very low (noise mediated) activity	no shmoo, since no signaling (and since Cdc42 is also not	chanaed. Explanation: Gpa1 is the only way how pheromone stimulation could reach downstream targets.	
					localized by Ras)		
		1		Absolutely no activity	No shmoos	Conclusion: There is indeed no signaling without Gpa1. It maybe necessary for the activation of a downstream component; if Gpa1 is not present, all the noise mediated activation is filtered out. A gpa1? & Ras.Val profile could confirm, that the noise generated by RasVal is indeed filtered out.	
09	scd1∆ (confirmat ion)	Starvation + Pheromone	h90	Same as in the 1st round. No / very low activity	Same as in the 1st round. No shmoos	Explanation: Same as in the 1st round.	
					No shmoos	Conclusion: Same as in the 1st round.	
10	Ras1.val17 & scd1∆	Starvation + Pheromone	h90	No prediction made.	No shmoos	No prediction was made.	
				Somewhat higher activity than scd1∆	No shmoos	Conclusion: The noise/ basal activity generated by RasVal is transcduced through the pathway	
11	Bvr1DD &	Starvation +	h90	Same as byr1DD:	No shmoos	Explanation: The Scd1&2-Cdc42-Pak1	
	scd1Δ	Pheromone		Delayed saturation, no	No Shinous	complex does not have targets more	
				downregulation.		downstream than Byr2.	
					No shmoos	Conclusion: The Scd1&2-Cdc42-Pak1	
						complex does not have targets more downstream than Bvr2.	
	Remark.	Remark: <b>Predictions regarding morphology</b> of single mutants are often no real predictions. They are predictions by the model, but their shmoo-morphology was known before.					
	Remarka	Stanuation cal	e for tr	riments has to be repeated in a strain lacking both recentors' genes			
	nemark2:	They do not express P-factor receptor in a sex-specific manner,					

but we cannot exclude that they still express it on a basal level.

 

 Abbreviations
 GeB
 Morphology Branch of the pheromone pathway consisting MoB
 Conducted experiments are in green

 MoB
 Genetic Branch of the pheromone pathway consisting the numerous downstream elements involved in actin cytoskel Falsified predictions are red

TABLE 3: DESIGNED EXPERIMENTS 2. FULL TABLE IS AVAILABLE IN THE SUPPLEMENT.

# APPROACHING THE FUNCTION OF GPA1

If one intends to draw finer conclusions about biological function, the experiments has to be possibly closest to *in vivo* situation and resolution has to be high. My collaborators conduct experiments in an almost unperturbed system opposed to most of the works done in the field (See: *Data generation*).

We were also able to clarify the role of Gpa1 to some extent, which is very valuable, as Gpa1 is the most obscure component of the pathway. Clarifying the role of Gpa1 is a complex as well important issue; therefore a whole chapter is devoted for it. See: *Possible Gpa1 effectors*. In summary, it seems most likely that Gpa1 has multiple effectors. It is very likely to activate both Byr2 and Ras1, either directly or indirectly.

'Conservation of Spatial Information' concept helps to constrain the possible set of GPa1 effectors

# Possible Gpa1 effectors

The role of Gpa1 is unclear in the literature. It is known that Gpa1 acts upstream of Byr2 [91], and it semms very likely to be upstream of Cdc42-Scd1-Scd2-Shk1 complex. We are planning to prove that with a "*Byr1.DD* & *gpa1* $\Delta$ " strain. This strain has a Gpa1 independent, constitutively active transcriptional branch, so if shmooing is deficient, Gpa1 is an effector of the morphology branch.

Some authors believe that Gpa1 signals to Byr2 independently of Ras1 [18], [12]; whereas some sources and clues suggest a Ras1 mediated Byr2 activation is how Gpa1 acts. (See: in the Quiagen pathway repository and the next chapter) Our result strongly suggests that both activation ways exist. Firstly we demonstrated that there is a weak signaling even in the absence of Ras1 that tells that Ras1 can be detached, although it is still necessary for strong and functional signaling. Secondly Ras1 must receive a spatially correct signal to localize Cdc42 correctly. Ras1 is the only known component capable of localizing Cdc42 to the locus of shmoo growth; therefore it must receive a spatially correct activation signal. Gpa1 is expected to transfer this activity, although we could not prove this interaction directly. I developed a concept to support it by indirect evidences: *The 'Conservation of Spatial Information' concept (CSI)*.

Additionally to our current knowledge, a planned *gpa1Δ*&*Byr1.DD* strain could support that Ras1 activation is Gpa1 mediated. *Byr1.DD* strains have pheromone independent

Spk1 activation (see: *Byr1.DD strain*), although they grow shmoos only if pheromone is around. If the morphology branch is activated independently of Gpa1, this strain should grow shmoos as *Byr1.DD* cells do: it would have an intact morphology branch and a functional transcriptional branch. If, however they do not grow shmoos, it would mean that Gpa1 is necessary for the morphological response and probably for the activation of Ras1.

# THE 'CONSERVATION OF SPATIAL INFORMATION' CONCEPT (CSI)

There are two answers given to pheromone stimulus: expressional change that requires temporal stimulus, and morphological change (shmooing) that requires a spatio-temporal stimulus.

Primarily, this concept states that **components transmitting signal to the morphology pathway should conserve the local aspect of the information.** The pheromone gradient encodes both kind of information. Since there is only one pheromone receptor known, the two responses utilize mutual initial steps (Figure 20). Similarly, Gpa1 seems to be the only activated substrate of the receptor, and it is a membrane- and receptor- associated protein. Effectors of Gpa1 are unknown.

On the other hand, it is known that cytoplasmic proteins diffuse generally ~100x faster than membrane associated proteins. [92] Internal work at the Klipp group showed a 1000x difference in diffusion rate between cytosolic and integral membrane proteins. (C. Diener and E.Klipp, unpublished) In this work, diffusion rates for 18 cerevisiae proteins were estimated by the Einstein-Stokes-relation and cytosol and membrane viscosity data. It is assumed in all projects of spatial modeling known to me that cytoplasmic proteins lose the spatial character of the signal because of the fast diffusion. A review: [93]. Therefore, **the second statement of the concept is that all pathway components transmitting spatial signal to the morphology branch should exclusively be membrane bound or membrane recruited.** 

# CONSEQUENCES OF CSI ON POSSIBLE EFFECTORS

If one aims to restrain the set of possible Gpa1 and Ras1 interactions on a sound basis, one could utilize the 'Conservation of Spatial Information' concept. The only assumption to be made is that no other unknown indispensable proteins are involved in pheromone response. Again, essential mating genes were established by mutagenic screens. These methods are robust in identifying essential genes. (Alejandro Colman-Lerner, personal communication) Even if unlikely that unknown essential genes are hiding, we cannot exclude it, and therefore it still remains an assumption.

# **Consequences on Gpa1**

Gpa1 is directly activated by the pheromone receptor, the only protein directly activated by pheromone. The spatial signal has to reach Ras1, the only protein localizing Cdc42. There are no proteins known that can activate Ras1 without Gpa1. Also, our experiments showed that Gpa1-knockdown completely abolishes pheromone signaling. Consequently Gpa1 is very likely tp signal to Ras1 (directly or indirectly) besides signaling to Byr2. This limits the possible interaction partners. Our latest measurements showed that Gpa1 must functionally activate Byr2 in a Ras1 independent fashion. Gpa1 can functionally activate Byr2 by direct activation, or indirectly, e.g. by binding and activating Ste4 or by colocalizing Byr2 and Shk1, as Ras1 does. If the CSI holds, Gpa1 cannot use cytoplasmic ways to activate Ras1 with a spatially correct signal.

# **Consequence on Ras1 and Gpa1**

As stated above, Ras1 is the only component that can localize the Cdc42-complex, which is responsible to reorganize the actin cytoskeleton, thereby initiate shmoo growth. Theoretically there are two ways possible for the spatially focused activation of Cdc42. First, it can be globally recruited from the cytosol to the membrane, and locally activated at the site of highest pheromone stimulus. The second way would be that Cdc42 is locally recruited, and either globally or locally activated after pheromone stimulation. We could refuse the 1<sup>st</sup> hypothesis, as it was found that Cdc42 is recruited locally in another Ras1 mediated process. [94] The association of Ras1 and the Cdc42-complex happens only when Ras1 is activated by GTP. [91] To localize Cdc42 correctly, Ras1 also has to be activated in a spatially focused way.

According to the concept of 'Conservation of Spatial Information' there must exist an exclusively membrane-bound or membrane-associated chain of activation between Gap1 and Ras1. Gpa1 can functionally activate Ras1 locally by direct activation, or by binding (localizing) any GEF for Ras1: Efc25 or Ste6. The *ste6* $\Delta$  experiment has highlighted that Ste6 is not the exclusive transmitter of the signal [72], as the knockout of Ste6 causes much less severe phenotype, than the knockout of its substrate, Ras1. See: Figure 13. Gpa1 possibly uses Ste6 and Efc25 both.

30/10/2012

#### INTEGRATION OF A FULL-SCALE MODEL

Finally I constructed a model that describes the pathway with all known transcriptionally regulated components. Other proteins, which do not change in quantity or in activity, are only implicitly present. Since their activity is constant during the pheromone response, they are not variables rather parameters. Pyp1 and Pmp1 are part of the Spk1 deactivation parameter; similarly Gap1 is represented in the Ras1 GTP hydrolysis parameter, whereas Efc25 is present in the activation parameter of Ras1.

Although the model describes all major players in the pathway, it still simplifies in many aspects. Most importantly it is not a spatial model; it does not address the four subcellular locations of the pheromone response (extracellular space, cell membrane, cytosol and nucleus) by assigning proteins to separate compartments. In further, the activation by Ste4 and the regulation of Ste11 are simplified. The concrete regulation of Ste11 is very complex, see: [12]; the model reflects only its activation by Spk1. Ste4 is speculated to dimerize Byr2, but this is not directly proven, and the effect of dimerization is completely unknown. Nevertheless Ste4 is shown to be necessary for pheromone signaling [88,95]; only this proven information is incorporated in the model: Ste4 activates its substrate, standing for activation in general.



FIGURE 19 THE HYPOTHESIZED WIRING BEFORE THE LAST ROUND OF EXPERIMENTS. WE HAD NO EVIDENCE FOR A RAS1-INDEPENDENT ROLE OF GPA1.

In summary, the full pathway model incorporates most validated direct interactions among the pathway components. Simplification was applied in cases where too much is unknown about the interactions. Even though, assumptions were necessary regarding Gpa1, Ste4 and Ras1, since available knowledge is insufficient to create a functional model. Therefore, the model bases on a hypothetical network structure, depicted in Figure 20. The model was modified to be consistent with the newest set of data, which has shown that Gpa1 has at least one Ras1-dependent and one Ras1-independent effector. These measurements became available only at the beginning of October; such late incorporation of data posed quite a challenge.

The final model consists of 29 proteins connected by 52 reactions. All reactions are depicted as simple mass action kinetics, but for most of the proteins a general degradation rate was assumed, therefore there are only 45 reaction parameters. The WT, *Ras.Val* and *Byr1.DD* are the focal strains of this study, since they represent the 3 known distinct classes in behavior. See *Biology* section.



A coherent hypothesis for the pombe pheromone pathway

FIGURE 20 A POSSIBLE HYPOTHESIS FOR THE PHEROMONE RESPONSE. THIS STRUCTURAL MODEL IS CONSISTENT WITH THE LATEST RESULTS. THE MODEL CONTAINS A SIMPLIFIED VERSION OF THIS SCHEME.

The model was used to reproduce the behavior of these three strains. In the experiments, the removal of nitrogen from the media marks the beginning of the Ste11 dependent transcription. Our findings about starvation effects are ignored, as long as they are not confirmed. In the model, receptor transcription and Ste11 activation are explicitly dependent on nitrogen-starvation signal. In the latest model this is also relaxed to sole Ste11 activation; see section: *Further development* and modeling supplement.

The simulation starts from a vegetative steady state that corresponds to protein values measured by my collaborators and by [62]. The data is in all cases population average data, therefore the model is no single cell model. There were 3 time course (TC) datasets for Ras.Val and 2-2 datasets for WT and Byr1.DD strains, each around 10 time points (see in data supplement). Since the most data was available for Ras.Val strain the model was fitted to this dataset. (Copasi can fit one mutant at a time) The Ras.Val dataset also seemed the most informative about system behavior as ground using it. The other two datasets were used for validation.

# Model - Mark I

After a series of model development phases, I arrived at a stat e where the model could reasonably describe the datasets, as shown in Figure 21. The summed squared error (SSE, also: objective value) was 1.044.



FIGURE 21: MODEL MARK I IS CAPABLE TO DESCRIBE RAS.VAL DATA NICELY AFTER PARAMETER ESTIMATION. REMARKABLE IS THE EARLY SIDE PEAK REPRODUCED BY THE MODEL.

The model reaches steady state after stimulation and remains within general physiological concentrations (below 100 uM). It is acceptable for signaling proteins, but many pombe proteins are expected to be lower, therefore the model was further developed.

Although the presented model achieves steady state, it would not be necessary. Consider that this is a model of a sub process in mating, an encapsulating process that changes the transcriptional profile completely. The biological process of pheromone response neither necessarily goes into steady state, since the outcome of mating anyway completely switches the whole genetic program. The tight regulation of proteins in later times would anyway make no sense. In this respect, the system is very different from typical metabolic systems.

### Side peak

Most surprising was the fact that the model reproduced the early side peak of activation. It appeared in almost all of the datasets, however this data feature is very subtle, therefore much unexpected to be captured by numerical optimization. (See: description of problem at Figure 29) Indeed, multiplex model fitting returns many parameter sets that describe the data almost equally well, however the side peak appears rarely among them.



FIGURE 22: SIMULATED SPK1 ACTIVITY IN MODEL MARK I. STARVATION WAS INITIATED AT 100 HOURS.

Looking at the data I speculated that the early activation can be either due to the subsequent stimulation by starvation and pheromone, or by an initial stimulation by pheromone, and by subsequent self sensitization. (see: *Models for subsequent activation peaks*) Since the decisive experiments performed have to be redesigned and repeated, both options are open. In the model, the early stimulation peak is due to pheromone stimulation, and the second peak appears by the self stimulation of the pathway. This is illustrated with a time course of corresponding normalized variables.



FIGURE 23: THE NORMALIZED CONCENTRATION TIME COURSE OF DIFFERENT PROTEINS SHOWS THE REASON FOR THE DOUBLE PEAK. THE FIRST PEAK IS CAUSED BY PHEROMONE INDUCED BYR1 ACTIVATION, WHEREAS THE 2ND PHOSPHO-SPK1 PEAK IS CAUSED BY SURPLUS TRANSCRIPTION OF SPK. THIS IS A SURPRISING REALIZATION OF THE "SELF-SENSITIZATION MODEL" ELABORATED IN THE CORRESPONDING CHAPTER: MODELS FOR SUBSEQUENT ACTIVATION PEAKS.

In the 1<sup>st</sup> stage of activation, the Spk1 activity (red) is driven by transient Byr1 activation (blue), and then although phospho-Byr1 levels fall back, a second activation happens in Spk1. This is caused by increased transcription of Spk1 (green). Finally the signal is downregulated by the delayed transcription of "Downregulator" (purple). This model suggests that the main peak in pheromone response is the result of a commitment step: the main peak is more a 'response' to the first small peak, than to pheromone.

# Parameter identifyability

Multiplex parameter estimations generally over ~5 magnitudes with random staring points were conducted to elucidate whether parameters, which describe the data equally well are unique for all models. The model is very complex; therefore *non-identifyability* is expected for most parameters. [96,97] Again, the aim of the full complexity pathway was not to generate a completely identified therefore predictive model. The idea was to check the consistency of the intuitive pathway hypothesis depicted in Figure 20 by dynamical modeling.

Most of the parameters indeed show a strong dispersion within physical constraint values. The dispersion of the parameters is visualized below. A limit in interpreting these results is the fact that manual inspection of the results often filters out 70% of the fitted models for violating some biological constraints. The explanation behind this high failure rate is that in parameter estimation not all of these behavior-expectations can be

applied as model boundaries. Nevertheless the few rules that are true for this complete set, also apply the subset of correct models.



FIGURE 24: HEAT MAP OF LOG TRANSFORMED PARAMETER ESTIMATES SHOW THAT MANY PARAMETERS (COLUMNS) ARE NOT EVEN CONSTRAINED TO VALUES ON THE MAGNITUDE SCALE.

The few constraints can be established about parameters is that Gpa1, Ste4 and Shk1 are very strong activators, and that Ras1-Byr and Ras1-(Scd1&2-Cdc42-Shk1) form very stabile complexes in almost all optimal fits. (See model supplement) This is consistent with our general biological notion about the interactions, however they were never shown.

A simple way to deal with parameter uncertainly is to reduce their numbers. One can e.g. assume that regulation does not happen on the level of protein degradation, and set all degradation rate to the same value. This was implemented indeed; however multiplex parameter fits could find no optimal fit if Gpa1, Sxa2, Mam2 and Pheromone degradation rates were also fixed to the same value. After a more elaborate analysis of parameter distribution one can identify *core predictions*, i.e. parameter values that always appear in optimal models.

Regarding the dynamics behind the double activation peaks, the model predicts, that the side peak is external stimulation driven, whereas the second is transcription driven. The whole pathways second activation is driven by self-sensitization not by separate starvation and pheromone stimuli. The pathway reaches full Ras1 activation transiently, and this fast *switch like* activation is the main driver of the delayed expression wave behind the big 2<sup>nd</sup> peak.

# Ras1 basal activity

It was speculated that the *Ras.Val* mutation causes earlier peaking since it causes a higher basal activity and therefore the system easier crosses the critical activation threshold. Once the system reaches an activation higher than the critical value, the autocatalytic system falls into active state "unstoppably", e.g. even if the stimulus is removed. The model reproduced a higher Ras1 basal activity. It was separately tested, if higher pre-stimulation Ras1GTP levels indeed cause earlier peaking. (See: 02\_Modeling\_Supplement\Full\_model\Earlier\_results)



FIGURE 25: DECREASED GAP1 ACTIVITY (*RAS.VAL* OR *GAP1* MUTATION) LEADS TO INCREASED STEADY STATE LEVELS OF RAS1GTP BEFORE STARVATIONAL INDUCTION. THIS TRANSIENT IS NOT PART OF THE PHEROMONE RESPONSE, IT ONLY SHOWS, HOW THE MODEL ARRIVES TO DIFFERENT VEGETATIVE STEADY STATES. THE HIGHER BASAL RAS1GTP LEVEL IS SPECULATED TO BE THE MAIN CAUSE FOR THE *EARLY SHIFT* OF THE TRANSIENT SPK1 ACTIVATION PEAK IN *RAS.VAL* AND *GAP1* MUTANTS. (FIGURE 5)

After the behavior of the model was analyzed, its predictive capacity was tested. Surprisingly, after appropriate changes, the model semi-quantitatively predicted the WT behavior and qualitatively predicted the *Byr1.DD* behavior. Fitting to one mutant dataset, than subsequently predicting the other two is a very satisfying result, given that most of the parameters remain unidentified.

# The Ras.Val mutation in the model

The *Ras.Val* -> WT back-mutation corresponds the restore of Rgs1-Gpa1 interaction, this means in the model that the Ras1's GTP hydrolysis rate increases with max. 1-2 magnitudes. The effect of increasing hydrolysis rate on the pheromone response is depicted in Figure 26. The simulated change very closely overlaps with the observed WT dataset. I found that a ~50x increase of the hydrolysis rate (to 4.298) gives the best overlap with the measured experimental data. This increase in hydrolytic capacity is realistic for the effect of Gap1. The sum of all squared errors for the WT model and data is 0.51.



FIGURE 26: PARAMETER SCAN OVER RAS GTP-HYDROLYSIS RATES, STARTING FROM 0.1 (RAS.VAL STRAIN) TO 10 IN MODEL MARK I. THE MODEL IS NOT ACCURATELY PARAMETERIZED AS A RESULT OF SCARCE DATA. BESIDE THE REMARKABLY HIGHER ACTIVITY AT LOW HYDROLYSIS RATE, ONE COULD OBSERVE THAT THE PEAK IS SATURATING IN THE BACKWARD (Y) DIRECTION, SO IT IS PREDICTED THAT FURTHER DECREASE OF RAS1 GTP HYDROLYSIS RATE WOULD NOT INCREASE THE INTENSITY OF SPK1 RESPONSE. BESIDE INTENSITY, THE PEAKING TIME IS ALSO AFFECTED BY THE HYDROLYSIS RATE.

# Byr1.DD

The *Byr1.DD* mutation is implemented in the model as the knockout of ppByr1 dephosphorylation. This is a very rough approximation (as it still receives input from Byr2), that is to be elaborated in further models. Simulating the model with this change produces simulation that is qualitatively, but not quantitatively reproduces the *Byr1.DD* dataset. The WT model was "knocked out" for the mentioned reaction and simulated; this is presented in Figure 27.



FIGURE 27: THE BYR1.DD MODEL-VARIANT QUALITATIVELY REPRODUCED THE SPK1 ACTIVITY IN BYR1.DD CELLS. THE PREDICTION IS CORRECT IN THAT PPSPK1 SATURATES (AND NOT ACCUMULTES) AND IT REACHES ITS MAXIMUM ACTIVITY SLOWER THAN WILD TYPE.

The saturating Spk1 value is magnitude bigger than observed, but is still similar to our observation in the following respects. 1. The Spk1 activity follows a saturating kinetics instead of an overshooting (or exploding) behavior. 2. Spk1 achieves its final activity slower than wild type. Further development focused on the *Byr1.DD* behavior, and reached a 100 smaller prediction. (See: *Model - Mark II*)

The model still had a number of drawbacks. First, though proteins remained within physiological limits, they are higher than expected by comparing them to known or estimated steady state concentrations (see: *Genome wide protein expression data*). Secondly, the vegetative steady state protein concentrations are used only as initial values but not in parameter estimation. Third, *Byr1.DD* prediction is inaccurate. These problems are corrected in the following model.

#### Model - Mark II

The further developed model corrected for the mentioned problems. The measured, but not the estimated vegetative protein concentrations were used for pre-starvation fitting of three transcriptionally regulated proteins. Measurements and estimates for transcriptionally not regulated proteins were used as initial concentration as in the previous model. Constraint on protein levels was tightened to 50 uM for secreted-, or unspecific-proteins and to 25 uM for signaling-proteins. Absolute Spk1 measurements showed that Spk1 concentration remains around 1 uM, this threshold was incorporated as well. Around 3\*20 parameter fits were conducted. Particle Swarm global optimization algorithm was used with 3000 generations. Only 3 out of the last 20 fits returned an objective value below 1.5 (next best fits OV is  $3.5^{12}$ ). Manual inspection found only one fitted model to behave in a sensible way. This parameter set was a big improvement from earlier models in that all signaling proteins remain in the uM range, in accordance to the data about Spk1, which shows a rise form ~0.1 uM to ~ 0.5 uM. Sxa2 is an exception, but this is an externally secreted protein, whose effective level has to cope with diffusion, therefore it is expected to be expressed magnitudes stronger. Still as shown in the earlier model, Sxa2 levels are very easy to parameterize.



FIGURE 28: SIMULATED BEHAVIOR OF ALL PROTEINS IN MODEL MARK II. SXA2 IS LEFT OUT FOR THE CONVINIENCE OF VISUALIZATION. IT SATURATES AROUND 160 UM STRONGLY DELAYED AT AROUND 200H.

The surprising result of the new fits was that none of the 3 numerically acceptable fits formed an early side peak in Spk1 activity. Compare Figure 22 to Figure 30. In any further development version I was unable to find a parameter set with an early side-peak. A reason for not finding that parameter regime again may be the fact that numerical optimization often has problems with capturing delicate data features. It often comes up with a qualitatively different result that is numerically more similar to the data, as exemplified in Figure 29.

 $<sup>^{12}</sup>$  OV worse than  ${\sim}1.8$  generally denotes a model that does not capture the shape of the dynamics at all



FIGURE 29: NUMERICAL OPTIMIZATION FAVORS FIT1, AS THE SUM OF SQUARED ERROR IS SMALLER THAN IN FIT 2. STILL FIT2 IS CLEARLY A BETTER QUALITATIVE REPRODUCTION OF THE DYNAMICS IN THE DATA.

#### **Increased predictability**

The new model also correctly describes the WT strain by introducing the same change as in Model Mark-I. With the optimal GTP hydrolysis parameter of 11.2 the model describes the WT data with a sum of squared error of 0.52. The change in behavior is shown in Figure 30.



FIGURE 30: SIMULATED DIFFERENCE IN PHOSPHO-SPK1 DYNAMICS BETWEEN WT (LOWER LINE) AND RAS.VAL MODELS (UPPER LINE). ALL MODEL PARAMETERS WERE FITTED TO THE RAS.VAL DATA, AND THEN THE MODEL WAS SIMULATED ONCE. A MODEL WAS SIMULATED A 2<sup>ND</sup> TIME WITH THE PARAMETER WITH THAT THE MODEL BEST DESCRIBES THE WT DATA.

Predicted *Byr1.DD* behavior was much closer to biological range, reaching saturation around 4 uM. Other tests showed that Spk1 saturation value linearly scales with the Spk1 activation parameter, therefore further increase in accuracy is expected to be easy in further model development. The timing of activation is nevertheless accurately described by the model variant. The relation normalized simulation values to data is

depicted in Figure 31. The un- normalized Spk1 time course is shown further down in Figure 32 (the lower curve, in blue).



FIGURE 31: NORMALIZED PHOSPHO-SPK1 LEVELS IN THE BYR1.DD MODEL. SPK1 SATURATES AT 4 UM, NORMALIZATION ONLY HIGHLIGHTS THAT THE PREDICTED TIMING OF SATURATION IS CORRECT. THE DIFFERENCE IN THE EARLY DYNAMICS IS EXPECTED TO BE CORRECTED ONCE THE CONCRETE EFFECT OF STARVATION ACTIVATION IS CLARIFIED. THIS MODEL WAS CREATED BY SETTING PPBYR1 DEPHOSPHORYLATION TO 0 IN THE WT MODEL; IT WAS NOT FITTED TO BYR1.DD DATASET.

One note on *Byr1.DD* protein: although it contains a constitutively active Byr1 protein, no one showed actually whether the engineered protein is indeed as effective as the WT phosphorylated protein. If it is smaller, than no wonder that model overestimates the observed activity.

# Issues

Model Mark II still has a few drawbacks. First of all Sxa2 levels are relatively high, although this is solely an impression, as Sxa2 levels are completely unknown. The Ras.Val -> WT back-mutation results in exploding Sxa2 levels, instead of saturating. Nevertheless this was easy to correct for. Sxa2 has the single role to degrade pheromone, and this function is defined by the product of *Sxa2 level \*activity*. So I just had to find a different parameter set, where lower Sxa2 values with higher activity exert the same effect as in the original model. This was successfully achieved by optimizing the corresponding four parameters (See modeling supplementary). Figure 32 shows, that now Sxa2 is not exploding up to 100x faster Ras1GTP hydrolysis rates (*Ras.Val ->* WT back-mutation) while ppSpk1 still behaves correctly, exactly as above.



FIGURE 32: DECREASING RAS1 GTP-HYDROLYSIS DOES NOT LEAD TO THE ACCUMULATION OF SXA2 OR PHOSPHO-SPK1 IN THE "SXA2 CORRECTED" VERSION OF MODEL MARK - II.

Another critical point is why Mam2 induction is implemented as explicitly starvation dependent. It is known that many proteins require starvation for their induction. In the model I chose only one protein to be explicitly depending on starvation, whereas others implicitly depend on it. Strikingly, almost none of the Ste11 dependent proteins could be measured (Table 1) although all of them gave a signal at mRNA level. This could suggest that pheromone pathway proteins are indeed not there before starvation. This also opens the question *translational* regulation of HMG-box induced genes. Mam2 was the lowest expressed mRNA, therefore was receptor chosen to represent the pathways suppression before starvation.

#### FURTHER DEVELOPMENT

In the next model I removed the mentioned explicit dependence, and the model was still able to describe data similarly well (SSE around 1.3) with a refitted parameter set. All protein concentrations remain within their corresponding limits, and I found nothing problematic during manual inspection. However the model does not predict the *Byr1.DD* behavior nor predicts the delayed activation upon WT *back-mutation*. It is still possible that a parameter set can be found; it is just very hard to find it in a 42 dimensional parameter space. However this model at its current stage only tells, that it is possible to describe the data with less assumptions regarding starvation induction.

Further work should focus to predict further mutants by parallel fitting to all available dataset. By fitting the model to WT, Ras.Val and Byr1.DD datasets in parallel one could achieve better identified parameters, and may be able to predict the effect of further mutations. Nevertheless, as long as pieces of qualitative information are missing from the pathway, quantitative enhancement of the model is less important.

# ANNOTATION

Proving the validity and reusability of the model, considerable emphasis was put on detailed annotation of proteins and their interactions. Annotation of reaction and proteins helps to check the validity of each bit of information in the model. The flexible and straightforward annotation possibilities in Copasi are incredible; they provided a great help for this process. In summary, the model is generally well annotated. Proteins are linked to UniProt entries and their discovery is referred to original papers on PubMed. Sometimes complementary notes point out some details about reactions, global quantities and variables. Reactions are also generally annotated with links to the publications (on PubMed) that has porven the given physical interaction. Some of the annotation is only accessible in the corresponding "Annotations" Excel file, as Copasi not yet supports links eg. to SMART protein interaction database.

# THE RAS1 COLOCALIZATION MODEL

In other organisms, Ras1 seems not only to localize but also to activate its effector. [61,98–101] In pombe however Ras1 binding does not activate Byr2. [52] The idea in this model is to show, that concentrating two otherwise also active proteins by colocalization can be sufficient to deliver a specific, signal triggered response. The concept is detailed below.

Assume two compartments, cytoplasm with a volume of 10 (arbitrary unit) and a membrane compartment with a volume of 1 (arbitrary unit). Assume two reactants *Substrate* and *Enzyme* that produce *Substrate* <sup>ACTIVE</sup> and *Enzyme* once they collide. Both have the concentration of 0.5 in the cytoplasm.

$$Substrate + Enzyme = Substrate^{ACTIVE} + Enzyme$$
(1.3.)

Assume that *Ras*<sup>ACTIVE</sup> localizes both reactants to the membrane compartment, thereby concentrating them locally.

$$Ras^{ACTIVE} + Substrate = Ras^{ACTIVE} + Substrate_{MEMBRANE}$$
 (1.4.)

and

$$Ras^{ACTIVE} + Enzyme = Ras^{ACTIVE} + Enzyme_{MEMBRANE}$$
(1.5.)

The active substrate is the output of this system. Now describe the pathways activity before and after to the activation of **Ras**. For simplicity, assume that **Ras** is instantly converted to **Ras**<sup>ACTIVE</sup> upon stimuli; this in turn instantly localizes all **Enzyme** and **Substrate** molecules to the small membrane compartment. A general ODE looks like the following:

$$d \frac{[Substrate \ ACTIVE]}{dt} = volume \ * \ [Substrate] \ * \ [Enzyme]$$

$$(1.6.)$$

$$* \ reaction_parameter$$

The equation can be simplified by setting the constant reaction parameter to 1. So the bimolecular reaction prior to activation looks:

$$d[Substrate \ ACTIVE]/dt = 10 * [0.5] * [0.5] * 1 = 2.5$$
 (1.7.)
is the initial rate of activation. After the activation, when substrate and enzyme are both completely re-localized:

$$d\frac{[Substrate \ ACTIVE]}{dt} = 1 * [5] * [5] * 1 = 25$$
(1.8.)

is the initial rate of activation. This means the rate of activation of this bimolecular reaction is one fold increased upon relocalization.

Now let's see what happens if there is a monomolecular deactivation of substrate. (In biological sense, it can be bimolecular when e.g. catalyzed by a homogenously distributed phosphatase). The rate here depends on the amount of active substrate, not the concentration! Let us see the simplest case: there is a pool of active substrate (no production as above) and no inactive substrate. Then:

$$d\frac{[Substrate]}{dt} = volume * [Substrate ^{ACTIVE}] * [Phosphatase] * reaction_parameter2 (1.9.)$$

Deactivation prior to activation, simplifying by setting the constant reaction parameter and the equally distributed phosphatase to 1 (also a parameter):

$$d\frac{[Substrate]}{dt} = 10 * [0.5] * [1] * 1 = 5$$
(1.10.)

is the initial rate of deactivation. Deactivation after the activation:

$$d\frac{[Substrate]}{dt} = 1 * [5] * [1] * 1 = 5$$
(1.11.)

is the initial rate of deactivation.

It shows that the deactivation, or a monomolecular reaction is not affected by relocalization! It means the speed of approaching the steady state is much higher. But is the steady state value itself is different? Surely, since an unchanged deactivation (Eq. 15) with a faster activation (Eq. 11 vs. 12) leads to an elevated steady state. This is the basic idea, how a colocalization step could switch the system to a more efficient state.

The aim was to build simplistic top-down model to check whether a simple colocalization (concentration) step of two otherwise unchanged components (Byr2 and Cdc42-Scd1-Scd2-Shk1 complex) can describe the observations. The aim was to

73 / 120

minimize the number of variables while describing the spatial role of Ras1 correctly. If the model can describe the data correctly, it can support the hypothesis that the only role of Ras1 is to colocalize two interacting components, without affecting their activity. The model principally encodes the Ras1 mediated crosstalk of the two main branches of the pathway, TRAB and MOB. See: Figure 17

### **IMPLEMENTATION**

A strongly reductionist version of the pathway was implemented as a two-compartment model in Copasi. A graphical interpretation of the model is presented in Figure 33.



Model 1:

FIGURE 33: REACTION DIAGRAM OF THE RAS-COLOCALIZATION MODEL.

The model consists of 11 species and 15 reactions, compared to the 29 species 52 reactions of the full-scale model. One could argue that this model is still moderately complex, and this is certainly true. The reason why it cannot be further simplified is the relatively complex data (rich Spk1 dynamics) and the number of variable-states and transport-reactions introduced by compartment modeling.

The model variables and parameters do not correspond to concrete proteins or their properties, rather to "pathway-chunks". These chunks include the "Ras1 and everything upstream", denoted by Ras; the "Cdc42-Scd1&2-Shk1 complex" named MOB; and "MAPK-cascade + Ste11" called TRAB. What happens is that COT is activated by Pheromone, which itself is induced upon a starvation signal. Active Ras1 colocalizes the cytoplasmic *MOB* and *TRAB* into the small membrane compartment thereby enhancing the reaction among them. The active substrate (*TRAB\_*active) is then transported back to the cytoplasm, where it triggers the downregulation. Cytoplasmic *TRAB\_*active is the pathway output, this is fitted to the Spk1\_pp dataset.

### **RESULTS AND DISCUSSION**

The main idea was to provide support for the hypothesis regarding the function of Ras by modeling. The purposeof the model was to reproduce the key features of transcriptional response to pheromone, not to build a complete and elaborate model of the system. Most importantly, the model reproduces the data, as shown in Figure 34. The result supports that colocalizing role of Ras1 is sufficient to describe the activity of the transcriptional branch.



FIGURE 34: SPK1 ACTIVITY (BLUE) FITTED TO THE FAKE DATASET (RED CROSSES) IN THE RAS COLOCALIZATION MODEL.



FIGURE 35: SIMULATION RESULTS OF THE RAS COLOCALIZATION MODEL ON A LONGER TIMESCALE. A: SPK1 ACTIVITY ONLY IN SEPARATE SIMULATIONS SHOWS THAT EARLIER OR LATER START OF STARVATION DOES NOT AFFECT PHOSPHO-SPK1 DYNAMICS, B: BEHAVIOR OF ALL PROTEINS WITH STARVATION AT 100H.

The compartmentalization is an additional layer of complexity, and since there are many vague parts regarding the full pathway structure, this concept is not directly implemented as a part of the full scale model. The mechanism is nevertheless present in the full scale model in a form of an approximation. It is assumed that the difference in volume is at least two magnitudes; if so; we can neglect the process in the cytoplasm (as being 100x-slower). This is a strong assumption, which should definitely be confirmed before we can achieve a consolidated picture of the pathway. In the full model, instead of treating compartments with respective volumes and the reaction rate separately, the chance of passive -> active state transition is lumped into one parameter. This parameter stands for the product of the *volume* \* (volume independent) enzyme activity. This simplifies the representation to mass action equation.

Parameter identifyability was tested with 250 Levenberg-Marquardt optimization runs initiated at random points in parameter space. In fits not 5% worse than the best fit, 4 parameters were completely identified (a single value), another 4 were well identified (barely differ more than a magnitude), whereas 7 were not identified, although their values were not spread over the whole parameter range. See modeling supplement for results.

### MODELS FOR SUBSEQUENT ACTIVATION PEAKS

At the beginning of the project it was unclear what starvation actually induces the system. It was known to induce transcription, but it was unknown whether it causes Spk1 activity. Spk1 activation could be direct or indirect, through the transcription of pathway-components or their regulators.

I noticed that the Spk1 activation profile contains an early side peak. It is rather small; therefore my collaborators considered it as a result of experimental noise. However I found it in almost all datasets. See the western blot curves before the main peak in: *Time course data and cell morphology*. Two-phase response of the pheromone response was never reported before, and we could not explain it with the initial concept of pathway activation: Figure 2. I tried to explain it with different hypotheses. Two, literature- and data-consistent hypotheses could be formulated. The hypotheses consider only the dual activation, not how it is downregulated. Later it was supported by the results of *Full Scale Model Mark I:* a single common downregulator is enough to repress both peaks at different times.

Simplified versions of the hypotheses were translated into ODE models. The models contained 9 variables, 10 reactions and 10 parameters; model files can be found in modeling supplementary. To test whether they can at least qualitatively capable of dual activation, they were fitted to a set of points representing an exaggerated dual activation. The reason why the models were not fitted to the real dataset is discussed at Figure 29 in section: *Model - Mark II*.

### MODEL 1: DIFFERENT ACTIVATORS MODEL

The most plausible idea for dual peaking is that the first peak is caused by starvation, and the second by pheromone. In this model N-starvation triggers M-factor production. M-factor stimulates P-cells who secrete P-factor in turn<sup>13</sup>. P-factor in turn stimulates the modeled M-cells. This delayed activation by P-factor may be responsible for the 2<sup>nd</sup> peak. The concept is schematically depicted in Figure 36. In the model though, it was depicted as a delayed self activation. The self activation is a reasonable simplification of reality, where two, otherwise mostly identical systems cross-activate each other.

<sup>&</sup>lt;sup>13</sup> Alternatively, P-cells secrete pheromone independently of M-cells, only initiated by starvation. This makes no difference for the decisive experiment.

It was believed in the pombe community that N-starvation only prepares the pathway for signaling by cell-cycles arrest at G1 and the transcription of components, but not by activating Spk1. p 291 in [12] and Kayoko Tanaka personal communications. The 2<sup>nd</sup> model implemented this idea.



FIGURE 36: ALTERNATIVE MODELS FOR THE INDUCTION OF SPK1. THE FIRST MODEL GENERATES ONLY ONE PEAK IN "STARVATION ONLY" EXPERIMENTS, WHEREAS THE 2<sup>ND</sup> HYPOTHESIS GENERATES ZERO. WE OBSERVED ONE ACTIVATION PEAK.

### MODEL 2: SELF-SENSITIZATION MODEL

The second hypothesis bases on the classical assumption in which starvation does not result in Spk1 activation. In this model a single activation is considered: Pheromone. In the model, it is switched from 0 to 1 (as starvation in the previous). This causes the first activation peak. The second round of activation happens by the sensitization of the PW, namely that a positive regulator is transcribed upon this initial PW activation, which in turn activate the transcriptional branch, just as pheromone does. Functionally, in this model the role of the 2<sup>nd</sup> peak of Spk1 activity is kind of *commitment of the cell* to the same input rather a stronger answer to a stronger external stimulus. In the previous model, the *commitment* for mating can be interpreted on the cell *culture* level and not on the level of the individual cell. Figure 36 panel B depicts one concrete molecular hypothesis about a self-sensitizing model. At least six other positive feedbacks exist in the pathway. The model is not specific about them.

### RESULTS AND VALIDATION TRIAL

The simplistic ODE implementation of the two models is practically identical as both models make use of a delayed positive feedback to generate the 2<sup>nd</sup> peak. Both models

could describe the observation. However biologically there is an important difference between them. In the second case, the feedback is within the cell, whereas in the first case the feedback is realized by the cross stimulation of opposite sex cells in the culture. This gives us the chance to interrupt this feedback in one model, but not in the other. By removing opposite sex cells from the culture, there would be no second activation peak in the 1<sup>st</sup> model and absolutely no stimulation in the 2<sup>nd</sup> model. My collaborators performed the experiment and indeed found a small activation at 4 hours <u>not</u> followed by the typical big peak around 7-9 hours. This result is consistent with the first model of different activators. A control experiment for the finding could be done with a strain with a non-functional Spk1 protein. If the 1<sup>st</sup> small peak is present, but the 2<sup>nd</sup> bigger is missing, than indeed subsequent activations cause the dual peaks.

Although it seemed that the used *h*- strain senses nitrogen starvation only, since M-cells only express P-factor receptor (along with M-factor production) after N-starvation. Some colleagues shared their concerns, that M-factor receptor [19] is still in the genome, and we cannot exclude the possibility that it is unspecifically expressed. If so, the cells would sense the M-factor, which may generate the basal signal. For this reason a receptor eliminated strain is created and the experiment is planned to be repeated.

Parameter identifyability was tested with 100 Levenberg-Marquardt optimization runs initiated at random points in parameter space. In fits not 5% worse than the best fit, 5 parameters were completely identified (a single value), another 4 were well identified (barely differ more than a magnitude) and 1 parameter was spread over a bit more than two magnitudes. *Remark*: Many other hypotheses are also possible, which are not detailed here, but these either turned out to contradict previous findings, or do not make sense biologically. For instance a starvation induced self-sensitization (and double peak) could be possible, but it would mean that Spk1 is primarily functions as a starvation sensor, which is not sensible in light of previous research. Our measurements also refused this possibility. Also one should be aware that the side-peaks may only appear in specific experimental setups, even if they are not artifacts. Cells may behave differently, because the proportion of mating cells can vary or stress can be different. Nevertheless the fact that such peak appears at some condition already gives some insight to the pheromone response.

### MAPK MODELS

...for the likely position of a negative feedback

### ALTERNATIVES IN NEGATIVE REGULATION

There are several known negative regulators known in the system: Sxa2, Rgs1, Gap1, Rad24, Rad25, Pmp1 and Pyp1. The data shows that the signal is downregulated after a transient peak (see: *Time course data and cell morphology*). To produce such an overshoot before downregulation, a delay in the negative feedback is necessary. On the time scale of several hours it is reasonable to assume that the downregulation happens somehow involving transcription.

### **Direct regulation**

Downregulation of a pathway can be direct, when a negative pathway regulator is transcribed, or a pathway component/positive regulator is repressed during pathway activity<sup>14</sup>. No transcriptional silencing of any positive regulator is discovered in the genome wide expression study [31], therefore we can relatively safely exclude this possibility. Likewise, Ste11 is found to work as a transcriptional enhancer.

### **Indirect regulation**

Alternatively, regulation could follow an indirect scheme. In this case the regulators' expressions do not change, but an enzyme is expressed or suppressed, which then post-transcriptionally *regulates a regulator*. Since I found no sign of any indirect regulation under transcriptional role, I concentrate on direct regulation. Still indirect regulation cannot be excluded from the possibilities, as new function of known proteins is often found.

### Transcriptional and translational regulation

In recent years it became a quite popular finding, that proteins are not only regulated transcriptionally, but also translationally. Although this is a possibility, I found that in pombe the best predictor of protein concentration is still mRNA level. Protein isoelectric point, protein mass, ribosome occupancy, relative poly(A) length, relative PolII occupancy, or mRNA half-life showed much worse correlation to protein levels. (See: *Genome wide protein expression data* and data supplement). It is exactly what others

<sup>&</sup>lt;sup>14</sup> Increased degradation of specific proteins is also a theoretical possibility, but there is no sign for such specific proteases.

found in mammalian cells. Therefore I acknowledge mRNA level as an acceptable estimate of protein counts, and disregard translational regulation.

### TRANSCRIPTIONALLY CONTROLLED REGULATORS

Using the online gene expression database of the Bähler Lab<sup>15</sup>, I found that only two negative regulators, Sxa2 and Rgs1 are upregulated. See also: [9,102]. These therefore fulfill the criteria to generate an overshoot (delayed downregulation). Sxa2 is a protease degrading pheromone, and Rgs1 is a "G-protein Activating Protein", it enhances the GTPase activity of Gpa1.

Collaborators previously tested the knockout of Gap1, and Sxa2 and both strains still undergo delayed downregulation. Sxa2 is a bit special, because it almost does not downregulate in the first 24 hours. However, in terms of late sustained activity, this is very different from *Byr1.DD* which does not downregulate even in 36 hour measurements. Morphology of the cells is also strikingly different, Sxa2 is very similar to *Ras.Val* macroscopic phenotype, and does not resemble *Byr1.DD* cells. Rgs1 knockout has still to be tested, but it is anticipated by experimentalists to yield similar results as Sxa2-knockouts as their morphology is very similar. [9,102] From these results it seems that the major regulatory feedback loop is still missing.

Negative regulation upstream of Ras is unlikely, since *downregulation deficient Ras.Val* strains show almost intact or even stronger downregulation. If the downregulation would happen upstream of *Ras.Val*, it would contradict its molecular phenotype, as *Ras.Val* molecule is thought to retain its activity much longer. This lessens the probability of classical downregulation by receptor internalization or desensitization. A complete proof could be measuring activity in a mutant strain, harboring a fully constitutively active Byr2. (Strain is being designed) My collaborators hypothesize that the downregulator may be a specific phosphatase of Byr1 or Byr2. I tried to support or object this idea by dynamical modeling, and found two opposite effects, which hamper drawing a definite conclusion. Nevertheless I share the increasingly recognized view that negative results are also important to increase understanding, therefore I present this part of my work.

<sup>&</sup>lt;sup>15</sup> http://www.bahlerlab.info/cgi-bin/SPGE/geexview

### MODELING A GENERAL MAPK MODULE WITH A DELAYED DOWNREGULATION



FIGURE 37: GENERAL MODEL MAP1-2-3K MODEL. THE MODEL SUPPORTED THE INITIAL GUESS THAT THE FURTHER DOWN ONE PROTEIN FROM THE DOWNREGULATOR IS, THE MORE BLURRED SIGNAL IT RECIEVES. THEREFORE STRONG TRANSIENT PEAKS ARE MORE LIKELY CLOSE DOWNSTREAM OF THE REGULATOR. NEVERTHELESS AN OPPOSITE EFFECT ALSO WORKS IN THE SYSTEM: THE LONGER THE FEEDBACK LOOP IS, THE MORE LIKELY AN OVERSHOOT IS. THEREFORE, WITH THIS SIMPLE MODEL I COULD NOT CONCLUDE WHICH IS THE LIKELY POINT OF ACTION OF THE -YET UNDISCOVERED- PHEROMONE PATHWAY DOWNREGULATOR. NAMING FOLLOWS THE POMBE PHEROMONE MAPK-CASCADE.

I created a general model of a MAPK module to see, how the overshoot depends on its downstream distance from the regulative step, but not on the length of the regulatory loop. The idea was that in a cascade of activators, each step blurs somewhat an initially sharp signal (not even considering stochasticity); this means the bigger the distance between the downregulator and the observation is, the less likely is to observe strong overshoots. *Less likely* here only means, that parameter range for the behavior is narrower; which does not necessarily mean that it is biologically less probable.

I implemented the cascade depicted in Figure 37, and found that an overshoot in Spk1 can be achieved only within a limited parameter range. This range of parameter was much wider for the upstream Byr1, the protein just below the downregulation. The model was tested for a wide range of parameters. In all cases the transient peak was more blurred for the more downstream species, but not always smaller. One should notice that in this simple system a direct ppSpk1 -> ppByr1 negative feedback abolishes overshoot completely, regardless of parameters. There is no time window in direct

feedbacks for overshoot. (See an exemplary scan in supplementary: "MAPK\_poz\_of\_phosphatase/ Length\_of\_regulatory\_loop/") This has the same reason why oscillations are impossible in two component systems. [103] Also, in case of the pombe pheromone pathway I am not aware of direct negative coupling in the MAPK module.



FIGURE 38: A REPRESENTATIVE EXAMPLE OF TYPICAL ACTIVITY PROFILES. GENERALLY SPEAKING, THE FURTHER THE PROTEIN FROM THE REGULATOR LIES, THE SMALLER OR BROADER OVERSHOOT IS OBSERVED (BLUE->PINK->YELLOW). GREEN DENOTES THE STIMULUS. IT IS SWITCHED FROM 0 TO 1 AFTER 5 MINUTES.

It was concluded that the distance from regulation has a strong negative effect on chance that a random parameter set results in an overshoot. Further characteristics of this effect, like dependence on loop size, are not investigated in this model.

#### **OPPOSITE EFFECT OF LOOP-LENGTH AND DISTANCE FROM POINT OF REGULATION**

In such a simple dynamical system, two forces shape the strength of the transient peak in opposing directions. Firstly, the longer the regulatory loop is, the more likely an overshoot is, since the time window between activation and downregulation is bigger, and so a wider range of parameters can produce an overshoot. Secondly, as we have shown above, the closer the protein is to the regulator of the pathway, the separate effect of activation and downregulation are less blurred; therefore it is more likely to observe an overshoot. A negative feedback on Byr1 instead of the upstream Byr2 brings the downregulation closer, but also shortens the regulatory loop. For this reason one cannot tell which protein is more likely to be target of negative regulation, given that we observed a strong transient activation peak at Spk1. Still this set of *in silico* experiments highlighted two generally important features of simple, negatively coupled systems. Nevertheless I do not claim that these are new findings, only that they helped the understanding some details of this problem. Surely many classic mathematical textbooks have examples showing e.g. the effect of feedback loop length, although I am not familiar with the classic literature of mathematical biology.

For model files, further images see, the Models Supplementary.

### GPA1-RGS1 MODEL

Generalization of an explanatory model for Rgs1-Gpa1 interaction

### The general explanation of GPA1 behavior

A very interesting study was published recently about the Gpa1-Rgs1 system. [104] It was shown that Rgs1 overexpression decreases signaling activity as expected, but surprisingly cells lacking Rgs1 show a decreased signal activity too. Although the observed behavior can easily be due to a completely different function of Rgs (e.g. a role related to its nuclear localization - see localization supplementary and appendix & [105]), I found it appealing that the current interactions can also be enough to describe these striking observations. It is fascinating to understand how a single protein can possibly act in opposing directions. Such behavior could also means that by tuning the Rgs1 level in an artificial setup, the system switches the behavior given to the same input. Such switch can be very useful in synthetic biology. Incorporation of Rgs1 into a constructed Gpa1-system could also promote a more distinct switch-like response to graded input.



FIGURE 39: HIGHER RGS1 LEVELS SUPPRESS SIGNALING AT HIGHER STIMULUS LEVELS. AT THE SAME TIME THEY ALLOW STRONGER MAXIMAL ACTIVITY IF THE STIMULUS IS STRONG ENOUGH. THE BEHAVIOR OF 3X RGS1 IS UNEXPLAINED IN THE ORIGINAL PAPER, AND IT IS NOT CLEAR WHICH DYNAMICAL RELATION COULD PRODUCE SUCH BEHAVIOR. REPRINTED FROM SMITH ET AL.; 2009 CELLULAR SIGNALING.

The main message of the observation is that level of stimulus or input and the level of Rgs1 shape the output non trivial. Figure 39 suggests the following: when the stimulus stimulus/Rgs1 ratio is low, the output is suppressed. When the stimulus/Rgs1 ratio increases the output steeply climbs up to the saturating output level. With increasing

Rgs-level, the switch point is shifted to the right, meaning more input is needed to achieve the critical input/Rgs1 quotient.

A second effect to see is that the saturating activity increases with Rgs1 levels. The 3x plasmid copy dataset (dark green) does not show this relation (it should show the highest saturation, possibly out of the graph). Biologically this saturation is unexplained, but shown to be reproduced by their model. Also notice, if the last data point is measured wrong, the fitted spine could also be very different than saturating. For what mechanical reason does it saturates low is not explained in the original paper, and since I was not provided with the model file upon request, I cannot tell it. My assumption is that this unexpected behavior arises as an effect of other parts of the rather complex original model. It means that this behavior rises from other reasons than the core Rgs-Gpa1 interaction. (There are only two reactions in their model catalyzed by Rgs1.) Nevertheless a simple re-implementation of the core of the model does yield a non-monotonic output behavior (increasing than decreasing), but cannot reproduce the '3X Rgs'- behavior.

The core of their model is depicted in Figure 40 Panel B. Increasing Rgs1 should always decrease the pool of Gpa1GTP\_INACTIVE. Reaction 2 & 5 both decrease this pool directly. The indirect positive coupling from reaction 5 (more substrate to go through r1 -> r3 -> r4) cannot be faster than the direct consummation of Gpa1GTP\_INACTIVE because of the mass conservation of Gpa1. Therefore the observed relation at 0x, 1x and 2x Rgs1-levels are possible to reproduce, but results at 3xRgs1 (Figure 39 above, dark green) cannot be explained *by principle*. How the published model done that is not clear to me.



FIGURE 40: DIFFERENT MODELS FOR THE GPA1-RGS1 SYSTEM. PANEL A: THIS SCHEME DEPICTS OUR ^GENERAL UNDERSTANDING OF A GPA-RGS SYSTEM. THIS SYSTEM IS INCAPABLE TO PRODUCE A NON-MONOTONIC RGS1-OUTPUT RELATION, AS EXPLAINED IN THE ANALYTICAL SECTION. PANEL B: "SMITH AND LADDS": RGS1 CATALYZES TWO REACTIONS. REACTION 2 ALWAYS DECREASE SIGNALING, BUT REACTION 5 CAN ENHANCE SIGNALING, IF THERE IS VERY STRONG STIMULATION, AND ALL GPA1 WILL EVENTUALLY BE STUCK AT THE GPA1GTP-INACTIVE POOL. PANEL C: NEW ALTERNATIVE MODEL 1. PANEL D: NEW ALTERNATIVE MODEL 2. THERE IS NO PRINCIPAL DIFFERENCE BETWEEN C AND D, THEY ARE JUST TWO CONCRETE REALIZATION OF A GENERAL CONCEPT.

The original finding provides an explanatory model, which was found by the authors to be the only model describing measurements in a manually generated set of candidate models. This model bases on a set of assumptions that I found very fragile. Therefore my aim was to relax these assumptions, thereby give a generalized description. Also I aimed to show that alternative models with fewer assumptions could also account for the data.

Firstly I looked at their rather complex model (17 reactions) and tried to identify, which idea gives the core of this behavior. I found that the following assumptions could be generalized for other models:

• Gpa1 can only activate one substrate per its on activation.

• Gpa1 has a general -but not necessarily standalone- inactive pool that is reached after the substrate is activated.

These were the most important differences from the default model of G-protein cycle See: Figure 40, Panel A.

### ANALYTIC CONFIRMATION OF THE MINIMUM NUMBER OF STATES OF GPA1

It was found in [104] by numerical simulations that the default Gpa1-Rgs1 mechanism is insufficient to describe the observed behavior. Here I show analytically why it is so.

First we have to look at the activation mechanism of G-proteins. They activate their substrates by binding, not by covalent modification. This means the substrate is active as long as the G-protein is bound.

$$\begin{aligned} \text{Inactive\_substrate} + & G\_protein\_GTP \\ & \rightarrow & Active\_substrate\_G\_protein\_GTP\_complex \\ & \rightarrow & Inactive\_substrate + & G\_protein\_GDP \end{aligned} \tag{1.12.}$$

The known behavior of GAP<sup>16</sup> proteins is that they catalyze a GTP-hydrolytic step by activating the G-protein's internal GTPase activity, here the 2<sup>nd</sup> reaction. The general activation scheme of G-protein is depicted in Figure 40 Panel A. [106] By transforming it to a set of ODE-s I show that this scheme is insufficient to describe a non monotonic relation between Rgs level and signaling output. I simply use hereafter 'Gpa1' for the GDP bound form!

System of differential equations:

$$d\frac{[Stimulant]}{dt} = constant: c1$$
(1.13.)

$$d\frac{[Rgs1]}{dt} = constant: Rgs1$$
(1.14.)

$$d\frac{[Substrate]}{dt} = Gpa1GTP - SubstrateFIRING * Rgs1 * k4$$

$$- Gpa1GTP * Substrate * k3$$
(1.15.)

<sup>&</sup>lt;sup>16</sup> Rgs1 is a G-protein activating protein (a *GAP*).

$$d\frac{[Gpa1]}{dt} = Gpa1GTP * Rgs1 * k2 + Gpa1GTP - SubstrateFIRING$$

$$* Rgs1 * k4$$
(1.16.)

$$d\frac{[Gpa1GTP]}{dt} = Gpa1 * c1 * k1 - Gpa1GTP * Rgs1 * k2$$

$$- Gpa1GTP * Substrate * k3$$
(1.17.)

$$d\frac{[Gpa1GTP - SubstrateFIRING]}{dt} =$$

$$Gpa1GTP * Substrate * k3 - Gpa1GTP - SubstrateFIRING * Rgs1 \qquad (1.18.)$$

$$* k4$$

Now see, how the focal variables behave if Rgs1 is increased  $0 \rightarrow \infty$ .

- Stimulant = NA
- Rgs1 =  $\rightarrow \infty$ ; increases by definition
- Gpa1 = ->>∞; monotonically increases, since Rgs1 only appears in the positive part of the equation.

In other words, Rgs1 catalyzed reactions decrease both other pools of Gpa1: Gpa1GTP and Gpa1GTP-SubstrateFIRING. So, if the above Gpa1 monotonically increases, the sum of Gpa1GTP and Gpa1GTP-SubstrateFIRING variables must decrease to fulfill the Gpa1 mass conservation relation: Gpa1 + Gpa1GTP + Gpa1GTP-SubstrateFIRING = a fixed pool; let us set it to 1.

• Gpa1GTP + Gpa1GTP-SubstrateFIRING = ->> 0; monotonic decrease.

We are interested in the firing complex, which produces the output signal of this system. Rgs catalyzes the disassembly of the active complex, but is it possible that it indirectly increases the pool of it? If the sum of (Gpa1GTP + Gpa1GTP-Substrate) decreases, we have the following possibilities depending of the choice of parameter set:

- Either Gpa1GTP-Substrate itself decreases, this means Rgs always decreases output activity,
- or Gpa1GTP decreases, but this variable stays in the only positive term in the equation of *d[Gpa1GTP-Substrate]/dt* and all other numbers stay constant, therefore Rgs1 decreases output activity again.

• Finally it may decrease both Gpa1GTP and Gpa1GTP-Substrate; this again states the Rgs1 has a negative role on signaling output.

So in systems where Gpa1 has only two states, a firing-GTP-state, and an inactive state, output and Rgs has a monotonic negative correlation: If Rgs is increased, the output necessarily decreases. Now the above conclusion is not necessarily true in more complex systems where Gpa1 can only activate a single substrate, and has fourth, inactive pool.

### The original model

The idea was proposed in [104]. The core of their proposal is depicted in Figure 40 Panel B. They found this scheme by trying different models and finally arriving to one that can describe the data. Their proposed model bases on a set of assumptions that were never proven, and I found them very peculiar in a biochemical sense. Assumptions implicitly or explicitly made by the authors:

- Rgs1 cannot deactivate the G-protein-Substrate complex by GTP hydrolysis; it can only act on unbound Gpa1. This is contrary to the generally known behavior of GAP's: they activate G-proteins internal GTPase activity to release the substrate.
- Gpa1 unbound to any partners have a 3<sup>rd</sup>, "inactive" state, which is independent of its GTP/GDP state. One can assume two GTP-bound conformation, where the 2<sup>nd</sup> inactive conformation is reached (in significant proportion) only after the release the substrate. Gpa1 is protein of moderate size (46.2 KDa), and there is no evidence for two stable conformational GTP-bound state.
- This inactive state can be rescued by simple GTP hydrolysis. For this reason, the inactive state cannot have covalent modification. (Surely, one can assume further proteins that very rapidly remove this covalent modification only after the GTP hydrolysis another layer of complication)
- Gpa1-Substrate disassembly is independent of Gpa1's GDP/GTP state. This is the other side of assumption 1, and it is similarly atypical.

### Alternative models

My idea was to show that none of these assumptions are necessary to account for the dual role of Rgs, except two core assumptions:

- Gpa1 can only activate one substrate per it's on activation.
- Gpa1 has an inactive pool that is reached after the substrate is activated.

Additionally the following assumption is made to show a concrete alternative mechanism that can describe the observations:

• Gpa1 is only pre-activating its substrate by changing its conformation, but a second step is necessary for a fully active/firing substrate.

This second activation can be by the binding of another effector (a complex formation), or it can also be that Gpa1-bound substrate is e.g. phosphorylated by itself or another enzyme. In mathematical terms these steps are very similar (assuming irreversible steps), and the 'complex-formation' alternative is implemented for simulation. In the new model Gpa1 is still trapped in an inactive form, but it is not a new standalone conformational change, rather Gpa1 is trapped in its own effector-complex. The model is depicted in Figure 40 Panel C &D.

Although the following statements are nothing different from the default behavior of GAP-s, for completeness, one must consider that the following assumptions are also made:

- Rgs1 can deactivate the G-protein-Substrate complex (default behavior)
- Gpa1-Substrate complex dissociates only/mostly by GTP hydrolysis of Gpa1.

### Results

I reproduced the original model in a simple form, and simulated it with different levels of Rgs1. (See the original model scheme in the supplementary of the publication: [104] also in modeling supplement) This simplified original model could reproduce the bimodal output behavior. The model is not parameterized with data; therefore concrete numerical values are omitted from description. Further details and simulation results are in: modeling supplement/Gpa1-Rgs1\_model.

### The simplified Smith and Ladds model

The substrate level can be treated as constant or variable in the simulation; both are tested in simulation, and did not change the outcome. Here for reasons of simplification we treat it as constant.



FIGURE 41: PARAMETER SCAN OVER DIFFERENT RGS1 LEVELS AT A CONSTANT STIMULUS. THE STEADY-STATE LEVEL OF FIRING COMPLEX INITIALLY INCREASES, BUT LATER DECREASES (IN LIGHT PURPLE), THOUGH THE MAXIMAL ACTIVITY (PEAK VALUE) MONOTONICALLY DECREASES FOR A SINGLE STIMULATION LEVEL (DEEP PURPLE).



Concentrations, Volumes, and Global Quantity Values

FIGURE 42: 10 SIMULATIONS OF OUTPUT DYNAMICS WITH INCREASING RGS1 LEVEL SHOW HOW THE DYNAMIC CHANGES. INITIALLY THE PEAK ACTIVITY REMAINS CONSTANT, AND THE STEADY STATE LEVEL OF ACTIVE EFFECTOR INCREASES. FURTHER INCREASE OF RGS1 DECREASES THE PEAK AND THE STEADY STAT VALUE IN PARALLEL.

### The new alternative model

The alternative model depicted in Panel C of Figure 40 was also implemented. The simulation of this model at constant stimulus level and different Rgs1 levels yielded a similar, but not exactly the same observation. The parameter scan highlighted that in

this setup the steady state is always reached without an overshoot at any levels of signaling. Overshoot can still be possible at other parameter sets.



Effect of different Rgs1 levels on pathway output

FIGURE 43: HISTOGRAM OF 1000 SIMULATIONS SHOW THAT SYSTEM OUTPUT INITIALLY INCREASES, LATER DECREASE WITH INCREASING RGS1 ACTIVITY/CONCENTRATION.



Time evolution of pathway output upon changeing Rgs1 levels

FIGURE 44: TIME COURSE OF 25 SIMULATIONS WITH INCREASING RGS1 LEVEL SHOWS SIMILAR BEHAVIOR TO THE ORIGINAL MODEL WITH THE EXCEPTION THAT NO TRANSIENT PEAK APPEARS BEFORE STEADY STATE.

### Discussion

An interesting hypothesis was proposed in an earlier work to explain the observed dual role of Rgs1 only considering Rgs1's known interaction with Gpa1, and not assuming completely new functions to it. I provided an alternative model for the Gpa1 activation cycle, that base upon a smaller set of assumptions than the original model.

One must be aware, that this alternative model is also just a hypothesis unproven, so one could argue that this is no improvement in understanding of pombe pheromone pathway. My sole point was to relax the original requirements of an interesting hypothesis, and highlight the core of this mechanism that produces non-monotonic behavior. The simplified core motif that I achieved may appear in other biological or artificial systems. Similarly this model provided an interesting insight, how G-proteins with an inactive pool could change the effect of GAP protein. The alternative model yielded further interesting result. Before modeling the system I was wondering whether the release of Gpa1 from the double activated substrate complex is necessary to reproduce behavior. I speculated no to be. It turned out that is necessary at all tested parameter sets.

### Cell polarization in pheromone signaling

Successful mating requires the robust selection of the single best partner. This requires cellular polarization. What mechanism underlies polarization events in biology is not clear, and many models are proposed. An overview of models classified by the mechanism of spatial focusing is published recently by [93].

Many of these models were shown to reproduce some but not all aspects of biological polarization; however most of the proposed mechanisms are not clearly demonstrated on a molecular level. Although one could argue, having numerous models one should focus on validating them, I find highlighting that the currently described biological system can –at least- partially be sufficient for polarization also has its validity. This section is merely a play with thoughts, but it may serve as a source of inspiration for later spatial modeling.

The idea is that the spatially asymmetric activation followed by a delayed global inhibition can be sufficient for shmooing site selection. This global inhibition elevates the response threshold so that only the highest point of excitation can pass it. This concept seems similar to the "Local Excitation Global Inhibition" or LEGI model proposed by [107], without knowing it in detail.

In concrete biological terms, pheromone stimulus both triggers genetic- and morphological changes in Pombe. The transcription of negative pathway regulators is induced as the *sum of all stimuli* over the cell surface. Sxa2 and Rgs1, two negative regulators are expressed and secreted or spread in cytoplasm globally. These than globally repress signaling.



## FIGURE 45: SIMPLIFIED SCHEME SHOWING HOW A HOMOGENOUS NEGATIVE REGULATOR CAN HELP SHMOOING LOCUS SELECTION BY BIOCHEMICALLY SUBTRACTING A CONSTANT VALUE FROM A GRADIENT OF STIMULUS.

In the hypothesized morphological response, there is just a certain pool of protein is available; indeed, none of the *morphology branch* components are transcribed in mating.

Typically it is assumed that polarization involves a step of commitment. After this, the cell executes the response whenever a certain activation threshold is passed.

Cdc42 seems to be heavily involved in polarization in cerevisiae. [108] In a biochemically detailed model of Cerevisiae bud site selection, Cdc42 organizes itself into more stable clusters in a self-enhancing manner. [109] In this hypothesis, a similar mechanism is assumed. As long as the pool of necessary proteins is distributed (or trapped) along the whole cell surface, the cell cannot commit for one shmooing position. After global negative regulators repress the answer over most of the cell surface, the required protein (e.g. Cdc42) becomes available in sufficient amount for the place of highest excitation. Six positive feedbacks exist for the pathway. These may encode commitment, at least for the transcriptional response. That commitment by self activation can correct for the decreased stimulation caused by negative regulators.

Whether such a model could work could be tested by an ODE-based compartment model. The concept how such model could work is depicted in the following picture.

Spatial regulation in S. pombe pheromone pathway



The working hypothesis is that the interplay of very slowly diffusing membrane components and fast diffusing cellular components can give rise to a very precise localisation

FIGURE 46: DETAILED CONCEPT FOR A COMPARTMENT MODEL TO TEST WHETHER KNOWN COMPONENTS OF POMBE PHEROMONE RESPONSE CAN HELP CELL POLARIZATION.

This time compartments describe adjacent areas on a simplified one dimensional cell surface. Each compartment receives a different input stimulus representing stimulation by a pheromone gradient. Since the diffusion in membrane is very slow compared to the cytosol (See in section *The 'Conservation of Spatial Information' concept (CSI)*), transport only implemented between a globally mixed cytoplasm and the local membrane "boxes". The combination of local stimulation and global repression may give rise to spatially focused morphological response. Since the global negative regulation is a transcriptional step, it can produce any level of amplification.

This hypothesis is not proven to any level; it is nothing more than a possibility. It is neither implemented in a mathematical model; therefore I do not know whether the mechanism can indeed work as it is speculated to. Still I find it useful to think about that only the already proven interactions in the pathway can give rise to cellular polarization, without assuming any further component or interaction.

### THE DISTANCE MEASUREMENT HYPOTHESIS

It appeared to me that cells with strongly defective genetic activity (*Byr1.DD*) are still capable to find and attach to their partners, whereas other cells with only slightly altered genetic activity (e.g.  $gap1\Delta$ ) absolutely cannot find their partners.



## FIGURE 47: COMPARISON OF "MATING" MORPHOLOGY OF BYR1DD AND RAS.VAL CELLS. © EMMA KELLSALL & DR. KAYOKO TANAKA

I found it also interesting why all strains deficient in *common trunk* negative regulators (*gap1* $\Delta$ , *Ras.Val*, *rgs1* $\Delta$  [40,102] , *sxa2* $\Delta$ ) show very similar, but characteristic faulty morphology. My impression was -although many regulators exist in the pathway- that fine tuning is not necessary for the transcriptional program, only for the morphological change. I formulated this idea in a testable hypothesis.

### Shmooing in pombe

Shmooing right is a very difficult thing. Cells can grow a single shmoo of very limited length. Therefore cells should:

- Find out the single best partner
- Find out the precise direction
- Find out its distance (only grow shmoo if the partners is close enough)

The best partner is probably chosen as the highest pheromone source, its direction is likely to be signaled through the gradient of pheromone. Timing shmoo growth is more difficult thing, if cells respond to too distant cell they grow a shmoo that can never reach the partner. If they respond only to very close cells, they may wait forever for such. Therefore, they have to *establish how far the partner is*.

In general, one would expect that cells cannot rely on absolute concentration values, because population density could vary hugely, so should vary the background pheromone level. However it was shown that the addition of pheromone after a certain threshold induces shmoo growth without a gradient. [7] Similar behavior was found in

cerevisiae. [110] The natural variance in background pheromone level is possibly magnitudes smaller than the difference between a close partner and the background, explaining the existence of an absolute threshold. At least such difference is suggested by the fact that pheromone cannot be isolated from a liquid culture. (According to Kayoko Tanaka's experiences at Masuki Yamamoto's lab).

### The hypothesis

This hypothesis claims that the fine tuning of the pathway activity is necessary to accurately time shmoo growth, but -within limits- not necessary for a functional transcriptional program. A consequence of this statement is that disturbing or removing negative regulators in the upstream *common trunk* destroys the cells capacity to correctly establish whether the partner is close enough.

The genetic and the morphological pathways are still coupled. This means there is no morphology answer without Spk1 activation, and in turn there is no Spk1 activation, if the morphology pathways activity is repressed upstream of Shk1. Consequently, a rough intervention into the pathway, i.e. knockout of any backbone component –at least functionally- abolishes both responses.

#### SUPPORT FOR THE HYPOTHESIS

## Shmoos correctly find each other and attach even with strongly deficient genetic activity.

*Byr1.DD* is a strain containing a significantly disturbed transcriptional program, whereas an *intact morphology pathway*. This strain show a significantly delayed, non-downregulated Spk1 activity induced solely by starvation. Even with a so distinct transcriptional pattern these cells are capable to correctly find their mating partners and attach to them. Shmoos are generally not grown into the 'void'. This suggests that a change in genetic background is necessary, but not needed to be precisely timed and regulated. (Pheromone is in all cases necessary for shmooing)

# Strains deficient in the *common trunk* grow shmoos that do not reach any partner

On the other hand, *all strains* that have a mutation in negative regulation of the *common trunk* ( $sxa2\Delta$ ,  $Rgs1\Delta$ ,  $Gap1\Delta$ ,  $ras1\Delta$ ) show a single type of morphology, the '*Ras.Val*-class' phenotype. These cells grow thick, long, linear shmoos, but they never reach a partner. It would still be possible that these mutations affect the *transcriptional branch*, and the

changed transcription affects the morphology in turn (see: Figure 17). However we could exclude this possibility with the generation of the *Ras.Val & Byr1.DD* strain. Since Spk1 activity is solely starvation induced in *Byr1.DD* cells (as it does not take any input from the upstream Byr2), we know that *Ras.Val* mutation only affects the *morphology branch* in this strain. So the observed morphology is only due to the *mutations direct effect on shmooing*. In fact we observed *Ras.Val* morphology, while –as expected- the Spk1 activity remained *Byr1.DD*. Concluding, *the abnormal shmoos are solely due to the deregulation of the morphology response*. It is known that *Ras.Val* is a hypersensitive strain, as Ras1 has a high basal activity. Still why do these cells never find their partners? Simply arguing a too active shmoo-building nor explains the mechanism, nor provides real insights. I found two possible *specific* reasons to explain this.

One explanation is that cells cannot establish the direction of the partner, and the shmoo is just grown in a random direction. Alternatively cells grow their shmoo (with decreased precision) into the direction of the steepest pheromone gradient, they only lose their sense of *distance*. In other words, cells with a deregulated *morphology branch* do not know anymore when to start growing a shmoo; on the other hand, they lost their ability to suppress response to weak stimuli. This means that cells more or less respond on a *first come, first served* basis; they start growing their shmoos into the direction of the first incoming pheromone signal (although some diminished threshold may still exist).

The observation that these cells have much longer and thicker shmoos, may have partially independent grounds. These mutants lost a negative regulator; therefore their response is stronger. Also, partially it can be that the shmoos grow longer and thicker, since the cell struggles to reach its partner. These strong shmoos therefore may not be a cause of mating incapability, rather a consequence of it, or a side product of the diminished response threshold. I do not claim whether *Ras.Val* strains are capable of other steps in mating, like cell fusion. The cells may have further deficiencies besides locating their partner. Though the concept bases on a set of assumptions, that are not too hard to test. Even if the hypothesis fails, it could be a good food for thoughts. I designed a series of experiments to find out some basic facts about the precise regulation of partner localization.

### A PLAN TO SELECT THE CORRECT HYPOTHESIS

Based on our current observations, we do not know whether *Ras.Val* cells indeed can sense the direction of the partner, or whether they really respond to smaller stimulation.

### Stimulus threshold

First, we have to find whether *Ras.Val*, or any other *common trunk* regulatory deficient strains indeed respond to weaker signals than the wild type cells.

**Decisive experiment:** Liquid -culture and Synthetic P-factor is used for the experiment. WT and *Ras.Val* h- cultures are grown in parallel and separate samples are treated with each different amounts of synthetic pheromone. At some threshold WT cells start to grow (undirected) shmoos. If *Ras.Val* cells grow shmoos at lower pheromone levels, they indeed have a deficiency in response threshold.

### Directed shmoo growth

Secondly one has to prove whether the shmoos show a directional preference, or they are grown to random directions.

• **Decisive experiment:** The working hypothesis is that *Ras.Val* cells can sense what direction the partner is. The aim of the experiment is to prove that *Ras.Val* cells shmooing direction is non-random, they are grown towards a suitable partner cell.

A plate culture, suitable for imaging should be created with loosely spotted cell from:  $h+Byr1.DD \& ras1\Delta$  cells and h- *Ras.Val* cells. For control an equivalent plate of  $h+1^7$ *Byr1.DD & ras1\Delta* cells and h-*WT* cells could serve. In both experiments, the role of Byr1.DD & ras1\Delta P-cells, is solely producing P-factor pheromone independently of anything else than starvation. The ras1 $\Delta$  mutation only helps to distinguish P-cells, since these cells are round, and do not shmoo.

Imaging (time lapse at best) both cultures during shmooing under light microscope could highlight whether both WT and Ras.Val cells grow shmoos to the direction of the closest h+ cells. A critical step is that the plate-culture should not be too dense, that the direction of the closest partner is obvious.

<sup>&</sup>lt;sup>17</sup> *h*+ denotes P-cells, *h*- M-cells.

### **Distance measurement**

If it turns out from earlier experiments that *Ras.Val* cells can sense the direction, one could proceed further. One should than prove that *Ras.Val* cells are indeed deficient in establishing the distance of the mating partner.

**Decisive experiment:** The working hypothesis is that *Ras.Val* cells cannot sense if a partner is too far, therefore they grow shmoos even if they cannot reach their partners. The aim is to prove that *Ras.Val* cells shmoo towards the partner cells, even if they are "out of range". Similarly to above, a plate culture, suitable for imaging should be created with loosely spotted cells of:

Sex	Mutation	Morphology	Expected phenotype
			Grow small shmoos only h+ cells which are within
h-	WT	Small shmoos	'shmooing-range'
		Long linear	Grow long shmoos to the closest partner, even if it is
h-	Ras.Val	shmoos	far away
	Byr1.DD &	No shmoos,	Produce pheromone upon starvation, no change in
h⊥	ras1A	round colls	mornhology

TABLE 4: EXPERIMENTAL SETUP TO TEST THE WHETHER RAS.VAL CELLS INDEED HAVE A DEFECT IN MEASURING THE PARTNERS DISTANCE.

A critical step is again that the cell suspension should be diluted enough that suitable couples are placed far from each other! To precisely distinguish strains, one h- strain should constantly express GFP. Normal light yields sufficient fluorescent emission to identify these cells. For confirmation, the culture could be separated into two experiments: <u>h- WT co-cultured with h+ Byr1.DD & ras14</u> and h- <u>Ras.Val & h+ co-</u> <u>cultured with Byr1.DD & ras14</u>.

### CONCLUSION AND OUTLOOK

I consider this project was successful from many perspectives. First of all the project contributed some small details to the pombe research community. Secondly, the project resulted in a better understanding of the pheromone response, leading to new experiments and new conclusions. Thirdly, I learned much about how to address scientific problems. Furthermore, I enjoyed being responsible for a project from the beginning to its end.

### **Overview**

This research contributed a set of smaller details. Most of them are small advances from the current state of knowledge. We provided a molecular level overview of the full pheromone response. Surprisingly, there was no such present before. There were very rough representations available, concentrating mostly on the MAPK cascade and Ste11. Alternatively there are some studies published about the interactions or regulation of one or two components. Up to now there was no attempt to incorporate all available knowledge in one consistent frame. We presented such a complete description, and also validated it with *in vivo* experiments on its most critical points.

### Applicability of formal methods

Another conclusion of my work was that although a wealth of formal framework exists, their utilization is often very limited, as they often require much better data than typically available. A framework is needed that could generate all possible hypotheses of state transition from a set of proteins, functional and physical interactions. In this sense it would be similar to rule based model definition, but it would use interaction information to hypothesize state transitions, and not use state transitions to hypothesize 'contextualized reactions'.

### Pathway induction

The observation that all datasets show an early side peak led to the formulation of testable hypotheses. One explained the dual peaking of phospho-Spk1 by a single induction followed by positive feedback, the other with subsequent activations by starvation and pheromone. The decisive experiment has been performed, and favored

the subsequent activation hypothesis. However the experiment has to be repeated to ensure the conclusion.

### Role of Ras1

It is thought in the research community that Ras1 is a kind of side activator of the pathway, [12], [18]. Our results shown that, although Ras1-independent activation exists, Ras1 is a backbone component of the pathway (Panel A in Figure 13). What the activator mechanism of Ras1 in is unknown. The proven function of pombe Ras1 is the binding, but not activating its target Byr2. There are no works to my knowledge that explain how this molecular function of Ras1 concretely leads to signal transduction. This work presents the first coherent, molecular level hypothesis supported by data-driven modeling that provides insights in the underlying dynamical system. Beside its relevance in pombe, this concept –once proven- may help understanding other systems. In general, to my very limited knowledge, the concrete molecular effects of Rashomologues are not clear either. In mammalian systems, it is currently believed, that a combination of activation and localization is the mechanism how Ras activates its effectors. [61,98–100]

#### Pathway crosstalk

Although it was known for a very long time that pheromone signal activates transcription and the actin reorganization in parallel, also the interaction between components were found [91], the crosstalk of the two pathways was never addressed before precisely. We have *in vivo* validated a proposed interaction between the Morphology- and the Genetic- Pathway, and subsequently used this knowledge to dissect what is the effect of this interplay on shmoo morphology and transcriptional activity. Based on this interplay we could interpret both genetic and morphological phenotype with concrete molecular explanations, in a series of knockout experiments. Earlier research on pheromone pathway was not seeking to interpret the observed morphology of mutant strains; descriptions often stayed with simple terms like "mating deficient" or "sterile".

### Quantitative and predictive model

Finally we developed a mathematical model that describes the three focal mutant strains, and upon fitting to one (*Ras.Val*), could predict accurately the behavior of WT cells, and qualitatively the behavior of the other strain (*Byr1.DD*). To derive further predictions despite parameter non-identifyability, *model ensemble* based methods could be a choice. In these, one works with a set of models that describe the data equally well with possibly different parameter sets. One looks for predictions that appear in all, or in large class of models and tries to validate these. This basically involves a *cluster analysis* of simulation results. Such a method was applied and proven to be useful in [111].

### Critical self reflection

One clear failure of the project is that we still have not identified the key player in downregulating the Spk1 signal. More attention should be drawn on that as a prominent regulator of dynamics. The model should also be further developed. Since the pathway reconstruction taken up most of the time, its development started out late. Not all the data is readily incorporated into the model. WT and Ras.Val absolute Spk1 concentrations are not yet used for model fitting. This should be incorporated.

### **Closing Words**

I hope that you enjoyed reading this work, and that you found some interesting ideas.

### Appendix

### Homology with cerevisiae and general names

in Pombe	in Cerevisiae	Function / protein type	References
Mam2/Ma			
р3	Ste2/Ste3	Pheromone receptor	Tanaka 1993
Gpa1,			http://www-
signal	Gpa1, negative signal		bcf.usc.edu/~forsburg/ge
conveyer	regulator	G-protein alpha	netable.html
Does not		G-protein beta-gamma	Dohlman et al. 1993 PMID
exist	Ste4 & Ste18, signal conveyer	complex	8415763
	Sst2 (N-terminal domain		
	shows high sequence	Regulator of G-protein	
Rgs1	similarity!)	signaling	Pereirea & Jones, 2001
			http://www-
			bcf.usc.edu/~forsburg/ge
Ras1	Ras1 & 2	Colocalization / ??	netable.html
		Ras activation by GDP-	Hughes & Yamamoto
Ste6	Cdc25 (Some domains)	>GTP exchange	1990
Ste4	Ste50-some extent	Unspecific_PW activator	Barr & Wigler 1996
		Morphology PW effector	
Shk1	STE20, and 2 others	/ Genetic PW activator	Tu & Wigler 1998
			http://www-
			bcf.usc.edu/~forsburg/ge
Byr2+	STE11	МАРЗК	netable.html
			http://www-
			bcf.usc.edu/~forsburg/ge
Byr1+	STE7	MAP2K	netable.html
			http://www-
			bcf.usc.edu/~forsburg/ge
Spk1	Fus3, Kss1	МАРК	netable.html
	No apparent S. cerevisiae		
	homolog, Ste11 [cer] is	Global transcription	
Ste11	different!	factor	[12]
Sxa2 / Sxa1	Bar1 / Sst2	Pheromone protease	Imai & Yamamoto, 1992

### ALTERNATIVE NAMES OF POMBE PROTEINS

Names that come up in literature

Pombe protein def. name	Alternative name	General name
Sty1	Spc1	SAP3(?)K
Shk1	Orb2, Shk1, Ste20	MAP4K
Byr1	Ste1	MAPK

### Sex-specific genes

M-cells	P-cells	Function
mam2	map3	Receptor
mfm1, 2, 3	map2	Pheromone / precursors

n.a.			map1			Sex-specific MADS-box
						transcription factor
mat-Mc	is	constantly	mat-Pc	is	constantly	HMG-box transcription factor
expressed,		<i>Mat-Pc</i> is	expressed,		<i>Mat-Mc</i> is	pair (act also in heterodimer)
expressed upon starvation.			expressed u	ipor	n starvation.	
mat-Mi	is	constantly	mat-Pi	is	constantly	Homeobox transcription factor
expressed,		<i>Mat-Pi</i> is	expressed,		<i>Mat-Mi</i> is	pair of Mei2 (act in
expressed upon starvation.			expressed u	ipor	n starvation.	heterodimer)
Sxa2			Sxa1			Pheromone protease
Mam1			???			ABC transporter secreting Phe.
??			Map1			MADS-box TF

Source: table 18.1. in chapter 18 of [12]. Genes involved in pheromone "maturation" are excluded, see [12] p 282

# Online resources used in the project *string*

This database not only provides an interaction database, but provides a very informative summary information on proteins, the type and proof of interactions, gives direct links to publications and other databases. It proved excellent for pathway reconstruction, and was the most often used database in this project.

Link: http://string-db.org/

### Bähler Lab's online gene expression viewer

This online database presents large scale expression data in five different experiments with adequate control and replicates. The 'Pheromone Response / Mating', the 'Meiosis/Sporulation', and the 'Cell Cycle' experiments nicely highlighted regulation among the pathway components, also the Ste11 dependence was measured with a positive and negative control. It proved most useful in discovering the genetic regulation in the system.

Link: http://www.bahlerlab.info/cgi-bin/SPGE/geexview

### GeneDB: Schizosaccharomyces pombe

This database summarizes all kind of information about pombe genes, and contains links to further resources.

Link: http://old.genedb.org/genedb/pombe/

Its newer variant is the PomBase (http://www.pombase.org/), however it proved much less useful, simply because of its user interface

### Bionumbers

Bionumbers is a manually curated source that collects published quantitative data with references. It was used to estimate boundaries on the concentration of signaling proteins, establish cell- and nuclear- size, and some other values.

Link: http://bionumbers.hms.harvard.edu/default.aspx

### SPD (S.pombe Postgenome Database): Localizome

This database rose up from the supplementary images of a large scale localization study // HT localization. GFP tagged proteins' localization was checked during development of the model.

Link: http://www.riken.jp/SPD/Img\_page/32\_iP/32F08\_Loc.html

### PombeNet: gene conversion table

This is a summary table of Pombe-Cerevisiae homolog proteins.

Link: http://www-bcf.usc.edu/~forsburg/genetable.html

### UniProt

The fundamental protein database was used for searching for protein homologues by its Blast function.

Link: http://www.uniprot.org/blast/
### SUMMARY OF MUTANT STRAINS

		Spk1 activity (very rough			Cell shape
Strain	Condition	schemes)	Evidence	Shmoo (observed)	(observed)
	Starvation	Small peak (N-stry			(0.000.000)
wт	only	induced)	O&D WB	no	Normal
	Starvation +				
	Pheromone	Normal activity	WB	normal	Normal
	Starvation				
Byr1.DD	only	no downregulation	Q&D WB	kiss but not fuse	Normal
-	Starvation +	-			
	Pheromone	no downregulation	WB	kiss but not fuse	Normal
	Starvation	Small peak (N-strv			
Ras.Val	only	induced)	Q&D WB	no	Normal
	Starvation +			elongated, no	
	Pheromone	Earlier and higher peak	WB	partner is touched	Normal
	Starvation	Small peak (N-strv			
Ras	only	induced) ???	predicted		Round (!)
	Starvation +	Small peak (N-strv			
	Pheromone	induced) ???	predicted		Round (!)
	Starvation	Small peak (N-strv			
Scd1	only	induced)	predicted	no (!)	Round (!)
	Starvation +				
	Pheromone	very low activity	Q&D WB	no (!)	Round (!)
c. 12	Starvation	Small peak (N-strv			
Sca2	Only	induced)	predicted	no (!)	Round (!)
	Starvation +	vor low activity	prodicted	no (1)	Dound (I)
	Chamication	very low activity	predicted	no (!)	Round (!)
Dur 2	Starvation	vory low activity	prodicted	20	Normal
Dyrz	Stanuation +		preulcieu	110	NOTITAL
		very low activity	nredicted	no	Normal
	Starvation		predicted		Normai
Bvr1	only	verv low activity	predicted	no	Normal
	Starvation +		1		
	Pheromone	very low activity	predicted	no	Normal
	Starvation	Small peak (N-strv	,		
Gap∆	only	induced)	predicted	по	Normal
	Starvation +			elongated, no	
	Pheromone	Earlier and higher peak	WB	partner is touched	Normal
	Starvation	Small peak (N-strv			
sxa2∆	only	induced)	predicted	по	
	Starvation +			elongated, no	
	Pheromone	Earlier and higher peak	WB	partner is touched	

#### TRANSCRIPTIONAL REGULATION According to Bähler's expression viewer

		No transcriptional regulation (appr. flat expression		
Transcription induced by Ste11		profile)		
Name	Role	Name	Role	
Pheromone (Sex				
spec.)	Stimulant	Ras1	<i>COT</i> component	
Sxa_ (Sex spec.)	COT component	Gap1	Negative regulator	
Receptor (Sex spec.)	COT component	byr2	TRAB component	
Gpa1	COT component	Byr1	TRAB component	
Rgs1	COT component	Cdc42	MOB component	
	COT/TRAB			
Ste6	component	Scd1	MOB component	
	Positive Regulator			
Ste4	(?)	Scd2	MOB component	
Spk1	TRAB component	Shk1	MOB/TRAB component	
Ste11 (self				
inducing!)	TF	efc25	MOB component	
Pat1	Negative regulator	Rad24	Negative regulator	
Matc (Sex spec.)	TF	Rad25	Negative regulator	
Mati (Sex spec.)	TF			

"\_" character is the placeholder for sex specific name variants!

## SUBCELLULAR LOCALIZATION OF PPW-PROTEINS

Localization for all proteins in vegetative state stands in the supplement! Source: [10]

Species	RIKEN database	Link	Comment	2 <sup>nd</sup> Reference
ras1	cytosol=nucleus:	http://www.riken.jp /SPD/43/43A06.html	Known example of wrong localization! (K.Tanaka, pers. communication)	
byr2	No Signal	http://www.riken.jp /SPD/33/33E12.html	periphery; cytosol	Ozoe'02
-	-		periphery; cytosol	Baumann & Albright'98
byr1	nucleus>=cytosol	http://www.riken.jp /SPD/15/15C10.html	Accepted: but no function is nucles is known	
rgs1	nucleus>>cytosol	http://www.riken.jp /SPD/23/23C06.html	ОК	Pereira and Jones 2001
sxa2	there is no signal on the images!!! They conclude: cytoplasm = nucleus	http://www.riken.jp /SPD/24/24G01.html	Not accepted, this protein i secreted	Ladds 2000; reviewed in: Didmon 2002
ste4	cytosol	http://www.riken.jp /SPD/13/13D07.html	ОК	

cdc42	nucleus>cytosol	http://www.riken.jp /SPD/50/50D07.html	-	
scd2	nucleus>cytosol; periphery at cell tip and site of septum formation	http://www.riken.jp /SPD/34/34H02.html	-	
gpa2	SPB; periphery at site of septum formation; nucleus>cytosol	http://www.riken.jp /SPD/23/23F07.html	-	
spk1	SPB; nucleus>cytosol	http://www.riken.jp /SPD/32/32E07.html	ОК	
sxa1	ER	http://www.riken.jp /SPD/25/25F10.html	-	
scd1	nucleus>>cytosol; periphery at site of septum formation; SPB?	http://www.riken.jp /SPD/39/39A01.html	-	
mam2	periphery at cell tip and site of septum formation; vacuole	http://www.riken.jp /SPD/15/15H10.html	ОК	
gpa1	periphery; cytosol	http://www.riken.jp /SPD/32/32F08.html	Likely to be coupled to Obara 1991 Receptor, since it is the same in Cerevisiae.	
efc25	cytosol	http://www.riken.jp /SPD/47/47F03.html	-	
ste11	nucleus>cytosol	http://www.riken.jp /SPD/33/33A02.html	ОК	
ste6	cytosol	http://www.riken.jp /SPD/39/39A11.html	ОК	
pak1	periphery at cell tip and site of septum formation; cytosol	http://www.riken.jp /SPD/39/39F04.html	ОК	

## LIST OF SUPPLEMENTARY MATERIAL

There are files which do not fit to these pages but could be of interest. These are listed here and provided as .zip archive. **The supplementary is accessible** from the author, and on the following website: http://vertesy.web.elte.hu/Pombe\_thesis/ *Data supplement*: All quantitative and some other high-throughput data used for modeling and pathway reconstruction is to be found in the data supplement. References mentioned in each folder. Western Blot data is confidential, as it is yet unpublished work of E. Kelsall and K.Tanaka. *Model supplement:* The model supplementary is a folder with model files, the datasets which were used in parameter estimation, Excel tables about analysis of the results and figures representing the model or showing some model behavior. The following models are presented: 01\_Full\_scale\_model; 02\_The\_Ras1\_colocalization\_modelFull\_model; 03\_Subsequent\_activation\_models; 4\_MAPK\_poz\_of\_neg\_regulator; 05\_Gpa1-Rgs1\_model. Annotation of reactions and species of the full model stands in: Annotations.xlsx.

# Bibliography

- 1. Bork P (2011) Consider context. Cell: 2011.
- 2. Sipiczki M (2000) Where does fission yeast sit on the tree of life. Genome Biol: 1–4.
- 3. Neiman a M, Stevenson BJ, Xu HP, Sprague GF, Herskowitz I, et al. (1993) Functional homology of protein kinases required for sexual differentiation in Schizosaccharomyces pombe and Saccharomyces cerevisiae suggests a conserved signal transduction module in eukaryotic organisms. Molecular biology of the cell 4: 107–120.
- 4. V. Wood R, Gwilliam, Rajandream M-A, Lyne M, Lyne R, et al. (2002) The genome sequence of Schizosaccharomyces pombe. Nature 47: 1215–1220.
- 5. Russell P, Nurse P (1986) Schizosaccharomyces pombe and Saccharomyces cerevisiae: a look at yeasts divided. Cell 45: 761–762.
- 6. Davey J (1992) Mating pheromones of the fission yeast Schizosaccharomyces pombe: purification and structural characterization of M-factor and isolation and analysis of two genes encoding the pheromone. The EMBO journal 11: 951–960.
- Imai Y, Yamamoto M (1994) The fission yeast mating pheromone P-factor: its molecular structure, gene structure, and ability to induce gene expression and G1 arrest in the mating partner. Genes & Development 8: 328–338. doi:10.1101/gad.8.3.328.
- 8. Kjaerulff S, Davey J, Nielsen O (1994) Analysis of the structural genes encoding Schizosaccharomyces pombe: identification of M-factor in the fission yeast a third gene, mfm3. Molecular and cellular biology. doi:10.1128/MCB.14.6.3895.Updated.
- 9. Imai Y, Yamamoto M (1992) Schizosaccharomyces pombe sxa1+ and sxa2+ encode putative proteases involved in the mating response. Molecular and cellular biology 12: 1827–1834.
- 10. Matsuyama A, Arai R, Yashiroda Y, Shirai A, Kamata A, et al. (2006) ORFeome cloning and global analysis of protein localization in the fission yeast Schizosaccharomyces pombe. Nature biotechnology 24: 841–847. doi:10.1038/nbt1222.

- 11. Ladds G, Hughes M, Davey J (1998) Extracellular degradation of agonists as an adaptive mechanism. Seminars in cell & developmental biology 9: 111–118. doi:10.1006/scdb.1997.0215.
- 12. Richard Egel (Editor) (2003) The Molecular Biology of Schizosaccharomyces pombe: Genetics, Genomics and Beyond. 1st ed. Springer. p.
- 13. Forsburg SL, Nurse P (1991) Cell cycle regulation in the yeasts Saccharomyces cerevisiae and Schizosaccharomyces pombe. Annual review of cell biology.
- 14. Mata J, Lyne R, Burns G, Bähler J (2002) The transcriptional program of meiosis and sporulation in fission yeast. Nature genetics 32: 143–147. doi:10.1038/ng951.
- 15. Isshiki T, Mochizuki N (1992) Characterization of a fission yeast gene, gpa2, that encodes a G alpha subunit involved in the monitoring of nutrition. Genes & ....
- 16. Chung KS, Won M, Lee SB, Jang YJ, Hoe KL, et al. (2001) Isolation of a novel gene from Schizosaccharomyces pombe: stm1+ encoding a seven-transmembrane loop protein that may couple with the heterotrimeric Galpha 2 protein, Gpa2. The Journal of biological chemistry 276: 40190–40201. doi:10.1074/jbc.M100341200.
- 17. Shiozaki K, Russell P (1996) Conjugation, meiosis, and the osmotic stress response are regulated by Spc1 kinase through Atf1 transcription factor in fission yeast. Genes & Development 10: 2276–2288. doi:10.1101/gad.10.18.2276.
- Otsubo Y, Yamamoto M (2012) Signaling pathways for fission yeast sexual differentiation at a glance. Journal of cell science 125: 2789–2793. doi:10.1242/jcs.094771.
- 19. Tanaka K, Davey J, Imai Y, Yamamoto M (1993) Schizosaccharomyces pombe map3+ encodes the putative M-factor receptor. Molecular and cellular biology 13: 80–88.
- Aono T, Yanait H, Miki F, Davey J, Shimodag C (1994) Mating Pheromone-induced Expression of the mat1 -Pm Gene of Schizosaccharomyces pombe: Identification of Signalling Components and Characterization of Upstream Controlling Elements. Yeast 1.
- 21. Davey J, Nielsen O (1994) Mutations in cyr1 and pat1 reveal pheromone-induced G 1 arrest in the fission yeast Schizosaccharomyces pombe. Current genetics: 105–112.
- 22. Yamamoto M (1996) Regulation of meiosis in fission yeast. Cell structure and function 21: 431–436.
- 23. Yamamoto M (1996) The molecular control mechanisms of meiosis in fission yeast. Trends in biochemical sciences 11.

- 24. Xue-Franzén Y, Kjaerulff S, Holmberg C, Wright A, Nielsen O (2006) Genomewide identification of pheromone-targeted transcription in fission yeast. BMC genomics 7: 303. doi:10.1186/1471-2164-7-303.
- 25. Nielsen O (1993) Signal transduction during mating and meiosis in S. pombe. Trends in cell biology 3.
- 26. Willer M, Hoffmann L, Styrkársdóttir U, Egel R, Davey J, et al. (1995) Two-step activation of meiosis by the mat1 locus in Schizosaccharomyces pombe. Molecular and cellular biology 15: 4964–4970.
- 27. Nielsen O, Egel R (1990) The pat1 protein kinase controls transcription of the matingtype genes in fission yeast. The EMBO journal 9: 1401–1406.
- 28. Kofahl B, Klipp E (2004) Modelling the dynamics of the yeast pheromone pathway. Yeast (Chichester, England) 21: 831–850. doi:10.1002/yea.1122.
- 29. Hoffman C (2005) Except in every detail: comparing and contrasting G-protein signaling in Saccharomyces cerevisiae and Schizosaccharomyces pombe. Eukaryotic cell. doi:10.1128/EC.4.3.495.
- 30. Hughes D a (1995) Control of signal transduction and morphogenesis by Ras. Seminars in cell biology 6: 89–94.
- 31. Mata J, Bähler J (2006) Global roles of Ste11p, cell type, and pheromone in the control of gene expression during early sexual differentiation in fission yeast. Proceedings of the National Academy of Sciences of the United States of America 103: 15517–15522. doi:10.1073/pnas.0603403103.
- 32. Bao S, Qyang Y, Yang P, Kim H, Du H, et al. (2001) The highly conserved protein methyltransferase, Skb1, is a mediator of hyperosmotic stress response in the fission yeast Schizosaccharomyces pombe. The Journal of biological chemistry 276: 14549–14552. doi:10.1074/jbc.C100096200.
- 33. Kim HW, Yang P, Qyang Y, Lai H, Du H, et al. (2001) Genetic and molecular characterization of Skb15, a highly conserved inhibitor of the fission yeast PAK, Shk1. Molecular cell 7: 1095–1101.
- 34. Kitamura K, Shimoda C (1991) Schizosaccharomyces pombe mam2 gene encodes a putative pheromone receptor which has a significant homology with the Saccharomyces cerevisiae Ste2 protein. The EMBO journal.
- 35. Obara T, Nakafuku M, Yamamoto M, Kaziro Y (1991) Isolation and characterization of a gene encoding a G-protein alpha subunit from Schizosaccharomyces pombe: involvement in mating and sporulation pathways. Proceedings of the ... 88: 5877–5881.

- 36. Fukui Y, Kozasa T, Kaziro Y, Takeda T, Yamamoto M (1986) Role of a ras homolog in the life cycle of Schizosaccharomyces pombe. Cell 44: 329–336.
- 37. Hughes D a, Fukui Y, Yamamoto M (1990) Homologous activators of ras in fission and budding yeast. Nature 344: 355–357. doi:10.1038/344355a0.
- 38. Hughes D a, Yabana N, Yamamoto M (1994) Transcriptional regulation of a Ras nucleotide-exchange factor gene by extracellular signals in fission yeast. Journal of cell science 107 (Pt 1: 3635–3642.
- 39. Tratner I, Fourticq-Esqueöute A, Tillit J, Baldacci G (1997) Cloning and characterization of the S. pombe gene efc25+, a new putative guanine nucleotide exchange factor. Gene.
- 40. Watson P, Davis K, Didmon M, Broad P, Davey J (1999) An RGS protein regulates the pheromone response in the fission yeast Schizosaccharomyces pombe. Molecular microbiology 33: 623–634.
- 41. Imai Y, Miyake S, Hughes D a, Yamamoto M (1991) Identification of a GTPaseactivating protein homolog in Schizosaccharomyces pombe. Molecular and cellular biology 11: 3088–3094.
- 42. Kjærulff S, Lautrup-larsen I, Truelsen S, Pedersen M, Nielsen O (2005) Constitutive activation of the fission yeast pheromone-responsive pathway induces ectopic meiosis and reveals ste11 as a mitogen-activated protein kinase target. Molecul and cellular biology and cellular biology. doi:10.1128/MCB.25.5.2045.
- 43. Asakawa H, Haraguchi T, Hiraoka Y (2007) Reconstruction of the kinetochore: a prelude to meiosis. Cell division 2: 17. doi:10.1186/1747-1028-2-17.
- 44. Li P, McLeod M (1996) Molecular mimicry in development: identification of ste11+ as a substrate and mei3+ as a pseudosubstrate inhibitor of ran1+ kinase. Cell 87: 869–880.
- 45. Kunitomo H, Higuchi T, Iino Y, Yamamoto M (2000) A zinc-finger protein, Rst2p, regulates transcription of the fission yeast ste11(+) gene, which encodes a pivotal transcription factor for sexual development. Molecular biology of the cell 11: 3205–3217.
- 46. Kitamura K, Katayama S, Dhut S, Sato M (2001) Phosphorylation of Mei2 and Ste11 by Pat1 kinase inhibits sexual differentiation via ubiquitin proteolysis and 14-3-3 protein in fission yeast. Developmental cell 1: 389–399.
- 47. Ozoe F, Kurokawa R, Kobayashi Y, Jeong HT, Tanaka K, et al. (2002) The 14-3-3 Proteins Rad24 and Rad25 Negatively Regulate Byr2 by Affecting Its Localization in Schizosaccharomyces pombe. 22: 7105–7119. doi:10.1128/MCB.22.20.7105.

- 48. Peng Z, Wang W, Schettino A, Leung B, McLeod M (2003) Inactivation of Ran1/Pat1 kinase bypasses the requirement for high-level expression of mei2 during fission yeast meiosis. Current genetics 43: 178–185. doi:10.1007/s00294-003-0384-5.
- 49. Kjaerulff S, Müller S, Jensen MR (2005) Alternative protein secretion: the Mam1 ABC transporter supports secretion of M-factor linked GFP in fission yeast. Biochemical and biophysical research communications 338: 1853–1859. doi:10.1016/j.bbrc.2005.10.156.
- 50. Qin J, Kang W, Leung B (2003) Ste11p, a High-Mobility-Group Box DNA-Binding Protein, Undergoes Pheromone- and Nutrient-Regulated Nuclear-Cytoplasmic Shuttling. Molecular and cellular. doi:10.1128/MCB.23.9.3253.
- 51. Chang EC, Barr M, Wang Y, Jung V, Xu HP, et al. (1994) Cooperative interaction of S. pombe proteins required for mating and morphogenesis. Cell 79: 131–141.
- 52. Tu H, Barr M, Dong DL, Wigler M (1997) Multiple regulatory domains on the Byr2 protein kinase. Molecular and cellular biology 17: 5876–5887.
- 53. Chang E, Bartholomeusz G (1999) Direct binding and in vivo regulation of the fission yeast p21-activated kinase shk1 by the SH3 domain protein scd2. ... and cellular biology.
- 54. Wheatley E, Rittinger K (2005) Interactions between Cdc42 and the scaffold protein Scd2: requirement of SH3 domains for GTPase binding. The Biochemical journal 388: 177–184. doi:10.1042/BJ20041838.
- 55. Hirota K, Tanaka K, Ohta K, Yamamoto M (2003) Gef1p and Scd1p , the Two GDP-GTP Exchange Factors for Cdc42p , Form a Ring Structure that Shrinks during Cytokinesis in Schizosaccharomyces pombe. 14: 3617–3627. doi:10.1091/mbc.E02.
- 56. Kelly FD, Nurse P (2011) Spatial control of Cdc42 activation determines cell width in fission yeast. Molecular biology of the cell 22: 3801–3811. doi:10.1091/mbc.E11-01-0057.
- 57. Tu H, Wigler M (1999) Genetic evidence for Pak1 autoinhibition and its release by Cdc42. Molecular and cellular biology 19.
- 58. Hopkins AL, Groom CR (2002) The druggable genome. Nature reviews Drug discovery 1: 727–730. doi:10.1038/nrd892.
- 59. Gibbs J, Sigal I, Poe M (1984) Intrinsic GTPase activity distinguishes normal and oncogenic ras p21 molecules. Proceedings of the 81: 5704–5708.
- 60. Fernández-Medarde A, Santos E (2011) Ras in cancer and developmental diseases. Genes & cancer.

- 61. Karnoub AE, Weinberg R a (2008) Ras oncogenes: split personalities. Nature reviews Molecular cell biology 9: 517–531. doi:10.1038/nrm2438.
- 62. S M, A S, S C, W C, R A, et al. (2012) Quantitative analysis of fission yeast transcriptomes and proteomes in proliferating and quiescent cells. Cell.
- 63. Onken B, Wiener H, Philips MR, Chang EC (2006) Compartmentalized signaling of Ras in fission yeast. Proceedings of the National Academy of Sciences of the United States of America 103: 9045–9050. doi:10.1073/pnas.0603318103.
- 64. Yamamoto A, Hiraoka Y (2003) Monopolar spindle attachment of sister chromatids is ensured by two distinct mechanisms at the first meiotic division in fission yeast. The EMBO journal 22: 2284–2296. doi:10.1093/emboj/cdg222.
- 65. Yamamoto TG, Chikashige Y, Ozoe F, Kawamukai M, Hiraoka Y (2004) Activation of the pheromone-responsive MAP kinase drives haploid cells to undergo ectopic meiosis with normal telomere clustering and sister chromatid segregation in fission yeast. Journal of cell science 117: 3875–3886. doi:10.1242/jcs.01248.
- 66. Chikashige Y, Kurokawa R (2004) Meiosis induced by inactivation of Pat1 kinase proceeds with aberrant nuclear positioning of centromeres in the fission yeast Schizosaccharomyces pombe. Genes to ...: 671–684. doi:10.1111/j.1365-2443.2004.00760.x.
- 67. Chenevert J (1994) Cell polarization directed by extracellular cues in yeast. Molecular biology of the cell 5: 1169–1175.
- Ladds G, Davey J (2000) Sxa2 is a serine carboxypeptidase that degrades extracellular P-factor in the fission yeast Schizosaccharomyces pombe. Molecular microbiology 36: 377–390.
- 69. Ghaemmaghami S, Huh W-KW, Bower K, Howson RW, Belle A, et al. (2003) Global analysis of protein expression in yeast. Nature 425: 737–741. doi:10.1038/nature02046.
- 70. Wu J-Q, Pollard TD (2005) Counting cytokinesis proteins globally and locally in fission yeast. Science (New York, NY) 310: 310–314. doi:10.1126/science.1113230.
- 71. Dohlman H, Thorner J (1991) Model systems for the study of seven-transmembranesegment receptors. Annual review of ....
- 72. Papadaki P, Pizon V (2002) Two ras pathways in fission yeast are differentially regulated by two ras guanine nucleotide exchange factors. Molecular and cellular .... doi:10.1128/MCB.22.13.4598.
- 73. Schwanhäusser B, Busse D, Li N, Dittmar G, Schuchhardt J, et al. (2011) Global quantification of mammalian gene expression control. Nature 473: 337–342. doi:10.1038/nature10098.

- Shahrezaei V, Swain PS (2008) Analytical distributions for stochastic gene expression. Proceedings of the National Academy of Sciences of the United States of America 105: 17256–17261. doi:10.1073/pnas.0803850105.
- Friedman N, Cai L, Xie X (2006) Linking Stochastic Dynamics to Population Distribution: An Analytical Framework of Gene Expression. Physical Review Letters 97: 1–4. doi:10.1103/PhysRevLett.97.168302.
- 76. Murphy K, Mian S (1999) Modelling gene expression data using dynamic Bayesian networks.
- 77. Perrin B, Ralaivola L, Mazurie A (2003) Gene networks inference using dynamic Bayesian networks. Bioinformatics.
- Sachs K, Perez O, Pe'er D, Lauffenburger D a, Nolan GP (2005) Causal protein-signaling networks derived from multiparameter single-cell data. Science (New York, NY) 308: 523–529. doi:10.1126/science.1105809.
- 79. Hlavacek WS, Faeder JR, Blinov ML, Posner RG, Hucka M, et al. (2006) Rules for modeling signal-transduction systems. Science's STKE : signal transduction knowledge environment 2006: re6. doi:10.1126/stke.3442006re6.
- 80. Tiger C-F, Krause F, Cedersund G, Palmér R, Klipp E, et al. (2012) A framework for mapping, visualisation and automatic model creation of signal-transduction networks. Molecular Systems Biology 8: 1–20. doi:10.1038/msb.2012.12.
- 81. Novère NL, Hucka M, Mi H (2009) The systems biology graphical notation. Nature ... 27: 735–742. doi:10.1038/nbt1558.
- 82. Jensen LJ, Saric J, Bork P (2006) Literature mining for the biologist: from information retrieval to biological discovery. Nature reviews Genetics 7: 119–129. doi:10.1038/nrg1768.
- 83. Strogatz SH (1994) Nonlinear dynamics and chaos: With applications to physics, biology, chemistry, and engineering. Westview Press. p.
- 84. Klipp E, Liebermeister W, Wierling C, Kowald A, Lehrach H (2009) Systems biology: a textbook. Vch Pub. p.
- 85. Kennedy J, Eberhart R (1995) Particle swarm optimization. Neural Networks, 1995 Proceedings, ... 4: 1942–1948. doi:10.1109/ICNN.1995.488968.
- Hoops S, Sahle S, Gauges R, Lee C, Pahle J, et al. (2006) COPASI--a COmplex PAthway SImulator. Bioinformatics (Oxford, England) 22: 3067–3074. doi:10.1093/bioinformatics/btl485.

- 87. Didmon M, Davis K, Watson P, Ladds G, Broad P, et al. (2002) Identifying regulators of pheromone signalling in the fission yeast Schizosaccharomyces pombe. Current genetics 41: 241–253. doi:10.1007/s00294-002-0301-3.
- 88. Barr MM, Tu H, Van Aelst L, Wigler M (1996) Identification of Ste4 as a potential regulator of Byr2 in the sexual response pathway of Schizosaccharomyces pombe. Molecular and cellular biology 16: 5597–5603.
- 89. Marcus S, Polverino a, Chang E, Robbins D, Cobb MH, et al. (1995) Shk1, a homolog of the Saccharomyces cerevisiae Ste20 and mammalian p65PAK protein kinases, is a component of a Ras/Cdc42 signaling module in the fission yeast Schizosaccharomyces pombe. Proceedings of the National Academy of Sciences of the United States of America 92: 6180–6184.
- 90. Kim DD-U, Hayles J, Wood V, Park H-OH-JH-M, Won M, et al. (2010) Analysis of a genome-wide set of gene deletions in the fission yeast Schizosaccharomyces pombe. Nature biotechnology 28: 617–623. doi:10.1038/nbt.1628.
- 91. Xu HP, White M, Marcus S, Wigler M (1994) Concerted action of RAS and G proteins in the sexual response pathways of Schizosaccharomyces pombe. Molecular and cellular biology 14: 50–58.
- 92. Moran U, Phillips R, Milo R (2010) SnapShot: key numbers in biology. Cell 141: 1262– 1262.e1. doi:10.1016/j.cell.2010.06.019.
- 93. Jilkine A, Edelstein-Keshet L (2011) A comparison of mathematical models for polarization of single eukaryotic cells in response to guided cues. PLoS computational biology 7: e1001121. doi:10.1371/journal.pcbi.1001121.
- 94. Peter M, Neiman a M, Park HO, van Lohuizen M, Herskowitz I (1996) Functional analysis of the interaction between the small GTP binding protein Cdc42 and the Ste20 protein kinase in yeast. The EMBO journal 15: 7046–7059.
- 95. Bauman P, Albright C (1998) Functional analysis of domains in the Byr2 kinase. Biochimie.
- 96. Raue a, Kreutz C, Maiwald T, Bachmann J, Schilling M, et al. (2009) Structural and practical identifiability analysis of partially observed dynamical models by exploiting the profile likelihood. Bioinformatics (Oxford, England) 25: 1923–1929. doi:10.1093/bioinformatics/btp358.
- 97. Raue a, Becker V, Klingmüller U, Timmer J (2010) Identifiability and observability analysis for experimental design in nonlinear dynamical models. Chaos (Woodbury, NY) 20: 045105. doi:10.1063/1.3528102.
- 98. Vojtek a B, Der CJ (1998) Increasing complexity of the Ras signaling pathway. The Journal of biological chemistry 273: 19925–19928.

- 99. Campbell SL, Khosravi-Far R, Rossman KL, Clark GJ, Der CJ (1998) Increasing complexity of Ras signaling. Oncogene 17: 1395–1413. doi:10.1038/sj.onc.1202174.
- 100. Takai Y, Sasaki T, Matozaki T (2001) Small GTP-binding proteins. Physiological reviews 81: 153–208.
- 101. Tamanoi F (2011) Ras signaling in yeast. Genes & cancer 2: 210–215. doi:10.1177/1947601911407322.
- 102. Pereira PS, Jones NC (2001) The RGS domain-containing fission yeast protein, Rgs1p, regulates pheromone signalling and is required for mating. Genes to cells : devoted to molecular & cellular mechanisms 6: 789–802.
- 103. Novák B, Tyson JJ (2008) Design principles of biochemical oscillators. Nature reviews Molecular cell biology 9: 981–991. doi:10.1038/nrm2530.
- 104. Smith B, Hill C, Godfrey EL, Rand D, van den Berg H, et al. (2009) Dual positive and negative regulation of GPCR signaling by GTP hydrolysis. Cellular signalling 21: 1151–1160. doi:10.1016/j.cellsig.2009.03.004.
- 105. Hill C, Goddard A, Davey J, Ladds G (2006) Investigating RGS proteins in yeast. Seminars in cell & developmental biology 17: 352–362. doi:10.1016/j.semcdb.2006.03.008.
- 106. Hamm H (1998) The many faces of G protein signaling. Journal of Biological Chemistry: 5–8.
- 107. Levchenko A, Iglesias P a (2002) Models of eukaryotic gradient sensing: application to chemotaxis of amoebae and neutrophils. Biophysical journal 82: 50–63. doi:10.1016/S0006-3495(02)75373-3.
- 108. Wedlich-Soldner R, Altschuler S (2003) Spontaneous cell polarization through actomyosin-based delivery of the Cdc42 GTPase. Science's STKE 299: 1231–1235.
- 109. Goryachev AB, Pokhilko AV (2006) Computational model explains high activity and rapid cycling of Rho GTPases within protein complexes. PLoS computational biology 2: e172. doi:10.1371/journal.pcbi.0020172.
- 110. Dorer R, Pryciak P, Hartwell L (1995) Saccharomyces cerevisiae cells execute a default pathway to select a mate in the absence of pheromone gradients. The Journal of cell biology 131: 845–861.
- 111. Brännmark C, Palmér R, Glad ST, Cedersund G, Strålfors P (2010) Supplementary: Mass and information feedbacks through receptor endocytosis govern insulin signaling as revealed using a parameter-free modeling framework. The Journal of biological chemistry 285: 20171–20179. doi:10.1074/jbc.M110.106849.