HUMBOLDT-UNIVERSITÄT ZU BERLIN

bachelor thesis

im Fach theoretischer Biophysik zum Erlangen des Abschlusses B.Sc. Biophysik an der Humboldt-Universität zu Berlin

Ein kombiniertes Modell für die Messung von extrazellulären Biomarkern mit Rückschluss auf die Aktivität des Hog-Pathways im Modelsystem saccharomyces cerevisiae

02. Juli 2018

Name: Jan Piotraschke (572797)

Gutachter: Prof. Dr. Dr. h.c. Edda Klipp

1 Abstract

Hier kommt meine englische Zusammenfassung

2 Zusammenfassung

und hier meine Deutsche

Contents

3 Introduction

3.1 system biology

System biology is there for the extraction of a system wide understanding of living organismen. This includes the interaction of multiple proteins, genes, metabolits et cetera, which are measured in the laboratory. This approach gets more significant in the analysis of executed omics experiments which easily results in data in the gigabyte range (Zitieren: Toward an integrated ...). Mathematical modeling has proven to be a promising tool for the study of the complex processes of environmental stress adaptation, to reveal the role of each biological component in the system and to reveal how the system level properties emerged from collected activites of individual components [1].

Currently limitations of the system biology approach are the usages and constructions of mathematical equations which should represent the biological system. This is a trade off between reduction of the system of interest without dimishing the quality of the information value or reasonableness intended digital twin. Another important problem is that there does not exists a complete biological understanding and knowledge of all system component. The system biological appproach is therefore only a heuristic approach (zitieren!!!)

In-depth insights of an investigated system are e.g. useful for medicine and the biotechnology sector (cite: Toward an integrated software platform for systems pharmacology!!!!) because this results in the improvement of well constructed mathematical models of a cell system could be useful for the design of target-oriented medications.

An increase in external osmolarity leads to a cell volume reduction. The cell counteract this high osmotic pressure by increased intracellular glycerol as an osmolyte and restores in this way its volume.

Mathematical models *in silico* are further helpful to test laboratory experiments *in silico* to identify meaningful experiments by construction of DoE (Design of Experiment). This helps to save the resources (e.g. money, time) of the experimentalist and could results in a deeper understanding of the underlying biological system.

3.2 Saccharomyces cerevisiae

The yeast Saccharomyces cerevisiae (*S. cerevisiae*) is a unicellular eucaryotic organism and belongs to the class of fungi [2]. It was the first eucaryotic organism where the whole genome had been full sequenced. [3]. *S. cerevisiae* is one of the beast characterized eukaryotic models [2].

In nature, the environment of S. cerevisiae varies in factors like temperatur, nutrient levels or osmolarity with the time and the cell must adapt with these changes. [4] The Hog-Pahtway in yeast has a significant role in the adaption process after an osmotic stress exposure. It normalize the volume of the cell and with that the water balance with an accumulation of the osmolyt glycerol inside, by closing the glycerol membran transporter Fps1 ([5], [6]) and the production of glycerol.

3.3 state of the art

It already exists multiple models for the hog pathway (signaling module), ion transport (transport module) and the volume regulation (volume module) (!!! alles hier noch mit Zitaten belegen). The signaling module keeps tracks of the stress response signaling pathways
Each of these models describe a part of the cell system while assuming other important aspects of the system as constant (see picture 1).

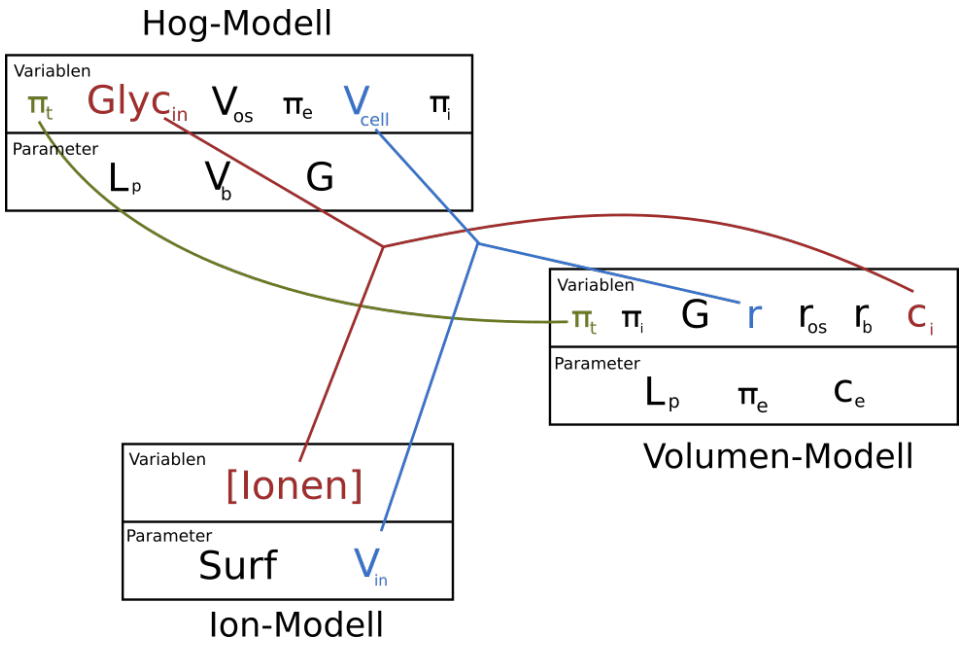

Figure 1: intersections of the three models

In our momentan state of knowledge, there is not yet a model which integrate this three modules into a single model. The combined model simulates the interaction between extra- and intracellular ionconcentration, changes in cell volume and the activity of the MAP cascade (???). The model senses the differences in the osmolarity between the cell and its environment and adapt with the important high osmolarity glycerol (HOG) signaling pathway with the MAP cascade the cell volume and the intracellular osmolyt concentrations.

3.4 theory of the used models

3.4.1 ion model

This model was designed to reproduces the experimentally observed potassium and proton fluxes indcues by the external stimuli KCl and glucose in S. cerevisiae. . It implemented eight well characterized transport proteins relevant to the regulation and maintenance of intracellular alkali-metal cation content. [7]

The ion model consists out of a non-equivilibrium thermodynamic (NET) approach to model the transport of the ion over the plasma membrane. NET is the analysis of spartial inhomogeneous systems and of time dependent processes. It is not valid if there are very fast processes or big inhomogeneity. The ion model assumes that both compartimens are well mixed. This way the NET approach hold true.

Per diffiniton, osmolarity means the amount of substance n of osmotic activ particles per volume V of the solution. Osmotic pressure accurs if two solutions with different osmolarity are seperated by a semipermeable membrane, where not all substances can pass the membrane for creating an equilibrium.
S. cerevisiae has a plasma membrane which functions as a semipermeable membrane. Water can diffuse freely through it, to adapt to the osmotic changes. At a hyperosmotic shock in the extracellular environment, water will flow out of the cell and let the cell shrink in volume. Cause of the shock event the high osmolarity glycerol (HOG) pathway in S. cerevisiae gets activ and synthesis the osmolyte glycerol. With that the cell increases the internal osmolarity and restores it's volume cause of the influx of water.

The immediate effect on yeast to an osmotic shock involves water outflow and decreasing volume [6].

Fluxes over membranes are irreversible processes. This will lead to a production of entropy in the system, which is calculated by the entropy production density σ which is respresented with the equation 1

$$\sigma = \vec{J_Q}\left(\frac{1}{T}\right) - \sum_{i=1}^{k} \vec{J_{c_i}} grad\left(\frac{\eta_i}{T}\right) + \sum_{r=1}^{R} J_r \frac{A_r}{T} \geq 0 \tag{1}$$

Equation 1 summeries the production of entropy under the conditions of temperatur difference, concentration difference and chemical reactions. Fundamentally, σ can only be created by fluxes J with it's forces X. In the area around equilibrium we can further assume that a flux J is linearly coupled over a phenomenological coefficient L with it's forces X, because at equilibrium all forces and fluxes vanish. Under this condition the following statement holds true.

$$\sigma = \sum_i J_i X_i = \sum_i \sum_j \frac{\delta J_i}{\delta X_j} X_j X_i = \sum_i \sum_j L_{ij} X_j X_i \tag{2}$$

L_{ij} were estimated from the experimental data of starved cells and are independent of the flows and forces [7]. They represents active transporters or channels in the plasma membrane. A high value of L_{ij} indicates that the transport of j is directly coupled to i. After some other assumption made from the ion model (for deeper insight see [7]), the equation for the flux of ion k is:

$$J_k = \sum_{j=1}^{n} L_{kj}(RT \cdot ln\left(\frac{c_k^{in}}{c_k^{out}}\right) + z_k F \Delta\phi) + L_{kAr} A_{Ar}$$

In the ion model a glycerol stimulus is further simulated. Ion regulation determines many physiological parameters, such as cell volume [1].

3.4.2 volume model

The volume model only holds true for a individual yeast cell in G1 phase. It can describes the small and steady volume variations during normal growth and the growth caused by hyper- or hypoosmotic shocks [8].
Turgor pressure π_t prevents exaggerated swelling and maintains cell shape [8]. π_t equals the hydrostatic pressure acting on the cell wall. For this the volume model desribes the elastically and plastically expansion of the membrane by involving the strain on the wall The volume model [8] describes further the water flux J_w between the cell and the environment with

$$J_w = -\frac{d}{dt} V_{os} = G * Lp * (\pi_e + \pi_t - \pi_i) \tag{3}$$

with Lp as the hydraulic conductivity, G the cell surface and the internal, external and turgor pressure (π_i, π_e , π_t). The cell volume V depends essential on the relation of the internal, external and turgor pressure due to changes in the osmotic volume V_os (water volume) because the model defined the total cell volume $V = V_{os} + V_b$ as the sum of V_{os} and a solid volume V_b which is not affected by water dynamics [8]. The internal and external pressure π depends on the concentration c of the osmotic active substance in the corresponding areas correlated over the equation 4

$$\pi = c \cdot R \cdot T \tag{4}$$

Furthermore, the model simulates the changes in the inner osmolyte concentration $\dot{c}_{in}$ with an uptake k_{uptake} and a consumption $k_{consumption}$ term:

$$\dot{c}_{in} = \frac{1}{V}(k_{uptake} G - k_{consumption} V - \dot{c}_{in}/V)$$

Summarising, it is highlighted that the cell expansion is mainly influeced by the interaction of

1. the control of the internal osmolarity c_{in}, which together with π_t drives J_w and therefore V_os

2. the elasto-plastic deformation of the cell wall due to π_t. $J_w < 0$ expands the cell wall and increase therefore π_t

3.4.3 hog model

The hog model is composed out of the Hog1 Mitogen Activated Protein Kinase (MAPK) cascade. This cascade is conserved even in higher eukaryotes including humans ([6]). The Hog1 MAPK is activated in the response to an increase in extracellular osmolarity [5]. All other non-osmotic stresses (e.g. temperature stres [5]) which are known to also activate the HOG pathway are not represented in this model. The MAPK pathways are important for transmitting and processing signals from the cell membran into the cell [6].
Hog1 MAPK is activated by the upstream Pbs2 MAP kinase kinase (MAPKK) by phosphorylation. The phoshorylated Hog1c (Hog1PPc) translocates into the nucleus where it activates different transcription factors by phosphorylation [4].

4 material and methods

4.1 software development

Cause a simulation for the combined models could take several minutes and the thereafter calculation and handling of this results are not garanted to work in the process of a program development, I came up with the idea that the results must be safed after each simulation and the analysis of them should be run in another program, as

first program : simulation –> save data
second program : data –> analysis

For the construction of this workflow I implemented a database and connected it to my python script. For the storing logic of the data I used a concept from CDISC for clinical trials which allowed me to construct the useful column names, data types and table in the database. Because a simulation of ODEs and algebraic equations can have pending sets of initial values, parameter values or equations terms, I expanded the concept of CDISC to the area of system biology and designed new data storage procedures. In picture (!!! Bild zeichnen lassen und hier einfügen) you can see the proposed workflow. The intention for the implementation of CDISC to system biology was the fact, that the FDA is moving towards CDISC standards for regulatory submissions [9].

With this approach, we can analysis and try new codes snippets with the results of a long simulation in seconds.
A version control mechanismen in the context of model construction was implemented with the help of a local database. The whole software development framework is visualized in (hier das Bild einfügen !!!!!!!!).

Furthermore, the storage of the model in a JSON file was in this way designed, that a natively typed equation in a .txt file format like

$$\frac{d}{dt}Na_{in} = \frac{J_{Na} \cdot G}{V_{os}} \cdot 10^6 - Na_{in} \cdot V_{ratio}$$

gets automatically seperated in the terms $\frac{J_{Na} \cdot G}{V_{os}} \cdot 10^6$ and $-Na_{in} \cdot V_{ratio}$. This allowes us to simulate each term component and analyse its constribution to the result of the equation in relation to the time.

5 Results

5.1 results of single models

Before merging the three models together it must be controlled whether each model is implemented correctly for itself. The pictures of the models simulation results are the guide for this. Sometimes there are discrepanies between the puplished model and the picture which should represent the model. Hereafter, the implementation process for each of the models is described.

5.1.1 Ion model

The challanges with the implementation of the ion model where, that the presented equation, initial values and parameters did not result in the intended system behaviour. After an in-depth analysis of the equation there where two anomalies:

1. the calculation of the change of the inner proton ion concentration has Bf as an undefined parameter
2. the fluxes have the wrong units

For solving the problem with the undefined parameter in (1) the ODE was constructed by deriving the formula for the calculation of the pH value for diluted solution

$$pH = -log_{10}([H^+])\frac{d}{dt}pH = -\frac{d}{dt}log_{10}([H^+]) = -\frac{1}{ln(10)}\frac{d}{dt}ln([H^+])$$

The published ion model has discrepancies in the parameter values of the phenomenological and stoichiometric coefficients and the calculation of the ion fluxes have the wrong units

The ion model assumes that the celllular compartment and the environment are well mixed. Their does not exists concentrations gradients.

5.2 hog model

The hog model consists of ODEs of the following form:

$$\frac{d[P]}{dt} = f(\vec{P}) - V_{ratio} \cdot [P] \tag{5}$$

Equation 5 describes in general with the first term at the right hand side the changes of a concentration due to a chemical process and with the second term the changes in the concentration due to changes in volume [1].

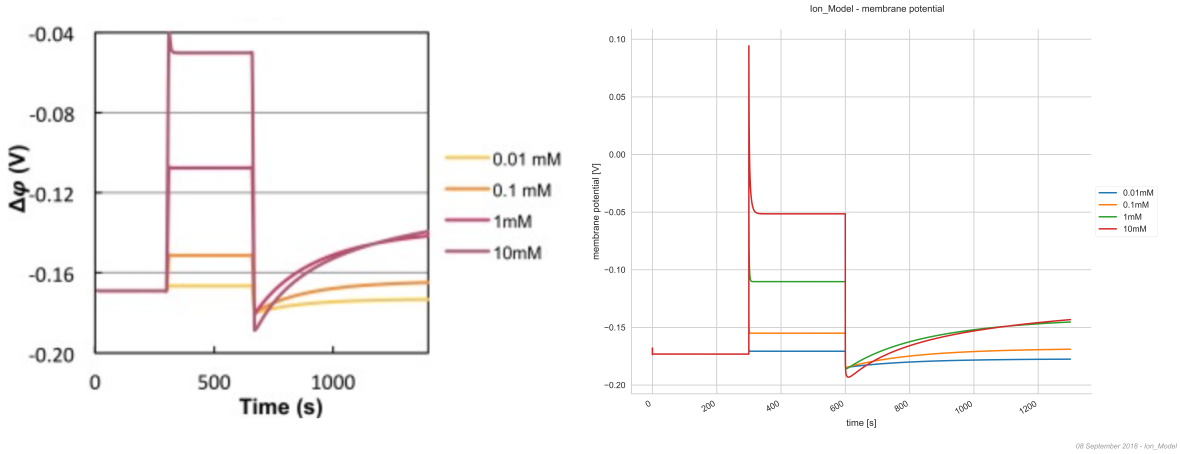

Figure 2: left: membrane voltage in the ion paper for different NaCl stimulus (0.01 mM - 10 mM); right: the simulated implemented ion model for the same stimulus range

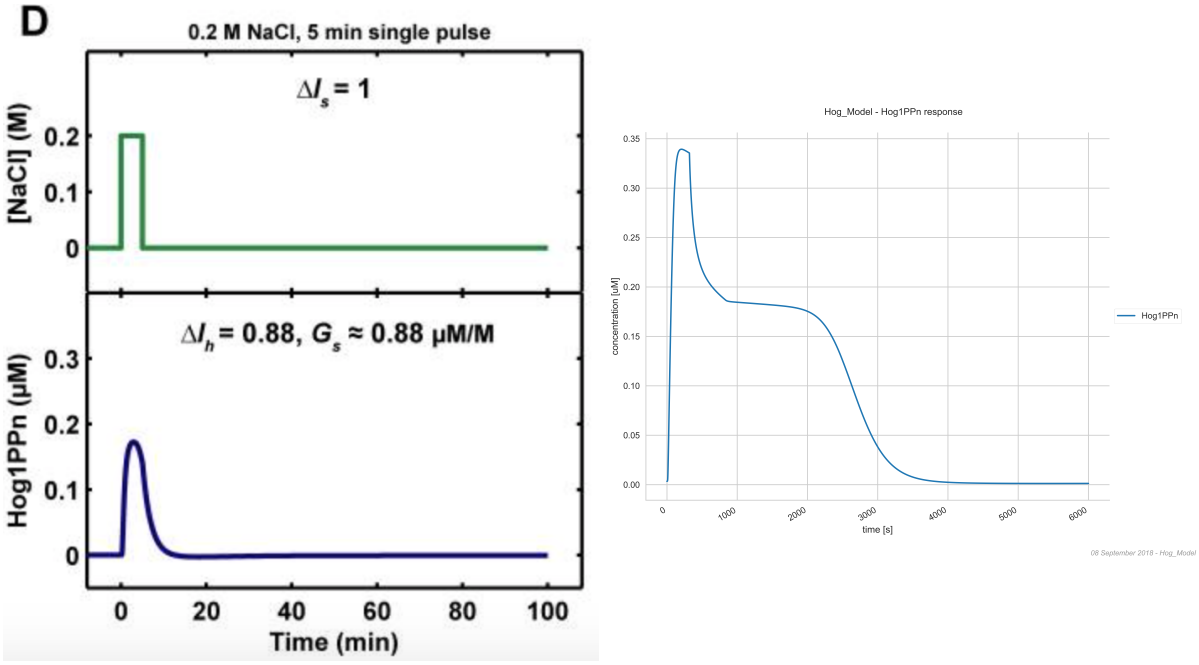

Figure 3: left below: membrane voltage in the ion paper for different NaCl stimulus (0.01 mM - 10 mM); right: the simulated implemented ion model for the same stimulus range

5.3 volume model

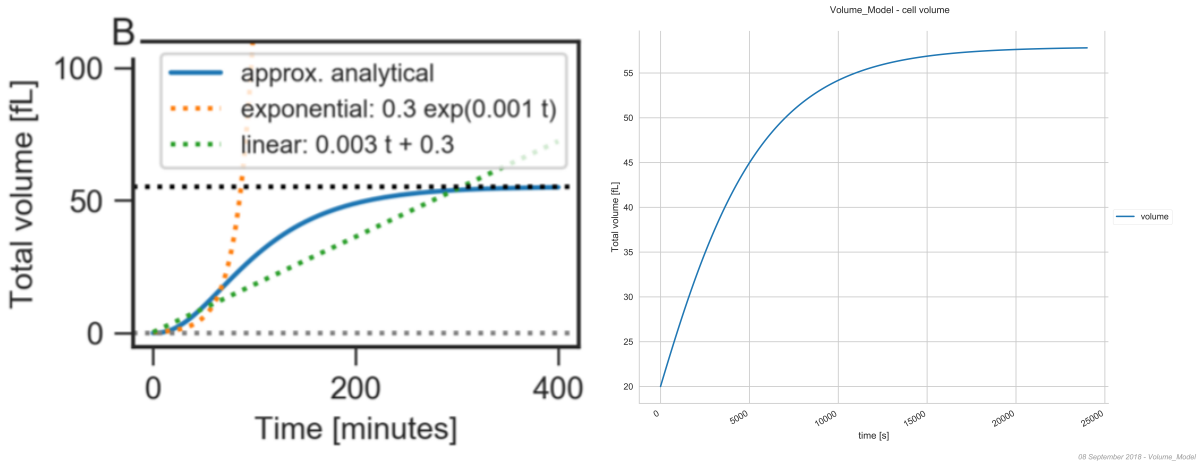

Figure 4: left: membrane voltage in the ion paper for different NaCl stimulus (0.01 mM - 10 mM); right: the simulated implemented ion model for the same stimulus range

5.4 merging of the models

The process of merging models is a tricky task. There exists some good guideline [10] for this. Hereafter, the used essential steps are sketched:

1. equations: One of the first steps in the process of model merging is the control whether there is a conflict between assumptions of the description of the biological systems. Conflicts must be resolved by combining equation terms or discard some informations

2. names: Find all the overlaps in the variable and parameter naming and convert them to a common sets of names

3. units: Standardize all used units for the parameter and initial values to a common set. In this process new factors where give to individual values if the unit changes requires that.

Changes in the equation design (ODE and algebraic equation) in the process of merging the model were done and are presented below:
In the whole process of merging we must consider their biochemical interpretation to maintain the plausblity of the model [10]. No parameter value was changed in his biological interpretation.

Table 1: changes in the equation design

	changes	reason
internal / external osmolyt concentration	$c = \sum([Ionen] + [otherosmolyte])$	simplification
internal / external presure	$\pi = c \cdot R \cdot T$	simplification
ion x changes	$\frac{J_x \cdot CellSurface}{CellVolume} \cdot 10^6$	unit harmonization
sorbitol stimulus	$\frac{d}{dt}[Sorbitol] = 0$	implementation of sorbitol s
uptake / dilution of osmotically active compounds	removed	replaced with the equation f

Initial values were changed because as we can see in 1 multiple components are used in several models. The components have different quantities in the initial assumptions. One of the main difference was the assumption of the cell volume V_cell in the hog model ($V_{cell} = 58fL$) and the volume model ($V_{cell} \approx 0.27fL$). To solve this problem we simulated the volume model until V_cell also have the value of $58fL$. We assumed the ODE values at this time point as the initial values for the corresponding substance in the merged model.

5.5 combined model

For the analysis and validation of the merged model we choosed the nuclear phophorylated Hog1 (Hog1PPn) as the control substance. Hog1PPn is also used in the hog model as the output because it regulates the expression of hundred of genes [11]. We borrowed the idea of a drug response curve from the theme field pharmocokinetic (zitieren: hier das Buch!!!) as a visualized output showed below.
(hier noch das falsche Bild, da Simulation noch nicht ganz fertig)

5.5.1 single models with the initial values from the combined model

After implement the initial values from the combined model back to the single models, no difference for the hog model was simulated. The reason for this lays in the simple reason, that we only changed the initial values from the ion and the volume model in the combined model.
The problem for compering the simulation results for the volume model is the fact, that the original data from the Github-Account of the ...

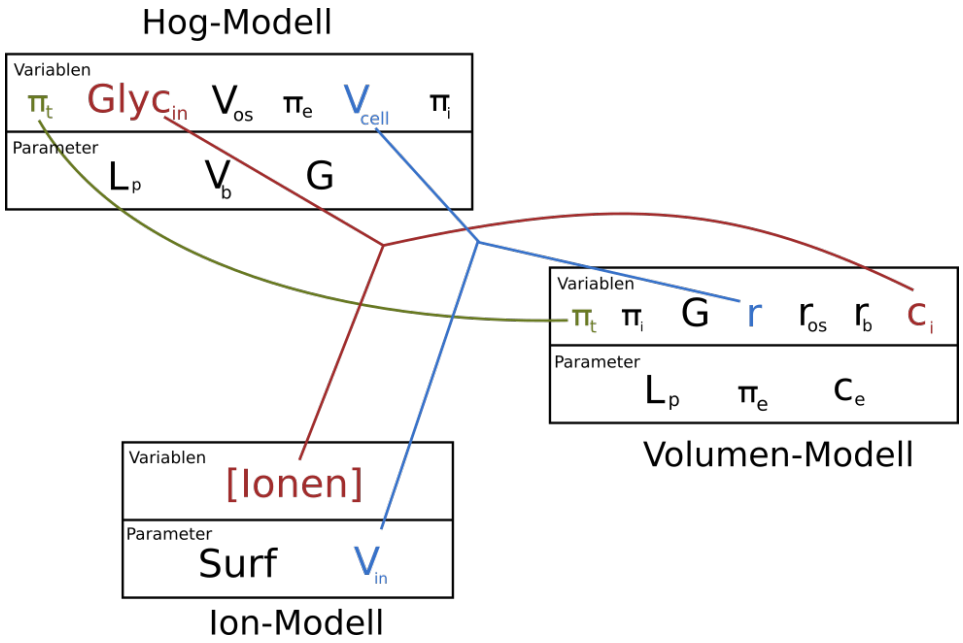

Figure 5: Response of Hog1PPn

6 Discussion

Distinctions between the 5.1.1 and 5.1.1 in the peak of the membrane voltage (???? nachschauen, ob es nicht eher membrane capacity ist und dann in der ganzen Arbeit entsprechende updaten) seems to have its root in the different time grid setting of the ODE-Solver in Copasi, which was used for the ion model simulation in the paper [7].
It is assumed that there are not any temperature gradient which would results in a heat flux.

Currently, only the proton and the potassium transport are ATP driven. It is known that Na^+ is also excluded by the Ena1p pump, which extrudes potassium. A general extension of the implemented membran transport processes could benifits the insight of then model in response of a external stimulus.

An implementation of the CWI pathways could further improve the information value of the combined model.

One of the problems in merging kinetic models from different sources and experimental conditions is the parameterization process of the merged model [12]. With no data available it was a try and error process to find a good parameter value for a substance, which is in a interplay with many variables. Available data sets could help in this process but even with this the parameterizations task remains a challenge. [1].

I propose that the idea to seperate an equation into their terms and storage them seperatly in a JSON file and that into a database, could lead to a machine learning based term knockout. This allows to reduces the computing time without limiting the equation result.

6.1 further work

I would like to expand the work with the model to the point where I am able to control plasma membrane transport mechanismen in eukaryotic single cells. I propose, that a precise anaylse of the extracellular environment could allow me then to get an idea what happend inside the cell to get this meassured secretion. For this secretory pathways (SEC) must be implemented to expand the stress-response network for the desired approach.
Another further work aspect is the improvement of the model vision control with the help of a database.

7 reference

[1] Ruian Ke, Piers J. Ingram, and Ken Haynes. An integrative model of ion regulation in yeast. *PLoS Computational Biology*, 9(1):e1002879, jan 2013.

[2] Serge Feyder, Johan-Owen De Craene, Séverine Bär, Dimitri Bertazzi, and Sylvie Friant. Membrane trafficking in the yeast saccharomyces cerevisiae model. *International Journal of Molecular Sciences*, 16(1):1509–1525, jan 2015.

[3] André Goffeau, Bart G Barrell, Howard Bussey, RW Davis, Bernard Dujon, Heinz Feldmann, Francis Galibert, JD Hoheisel, Cr Jacq, Michael Johnston, et al. Life with 6000 genes. *Science*, 274(5287):546–567, 1996.

[4] Jannis Uhlendorf. *Real-time feedback control of gene expression*. PhD thesis, Université Paris-Diberot, 2013.

[5] H. Saito and F. Posas. Response to hyperosmotic stress. *Genetics*, 192(2):289–318, oct 2012.

[6] Peter Gennemark, Bodil Nordlander, Stefan Hohmann, and Dag Wedelin. A simple mathematical model of adaptation to high osmolarity in yeast. *In Silio Biology*, 2006.

[7] Susanne Gerber, Martina Fröhlich, Hella Lichtenberg-Fraté, Sergey Shabala, Lana Shabala, and Edda Klipp. A thermodynamic model of monovalent cation homeostasis in the yeast saccharomyces cerevisiae. *PLOS Computational Biology*, 12(1):e1004703, jan 2016.

[8] Tom Altenburg, Björn Goldenbogen, Jannis Uhlendorf, and Edda Klipp. Osmolyte homeostasis controls single cell growth rate and maximum cell size of saccharomyces cerevisiae. unpublished, April 2018.

[9] Carey Smoak, Smitha Krishnamurthy, Mansi Singh, and Sy Truong. Forging new sdtm standards for in-vitro diagnostic (ivd) devices: A use-case, 2014.

[10] Wolfram Liebermeister. Validity and combination of biochemical models. 2008.

[11] Zhike Zi, Wolfram Liebermeister, and Edda Klipp. A quantitative study of the hog1 MAPK response to fluctuating osmotic stress in saccharomyces cerevisiae. *PLoS ONE*, 5(3):e9522, mar 2010.

[12] Zhuo Wang, Samuel A. Danziger, Benjamin D. Heavner, Shuyi Ma, Jennifer J. Smith, Song Li, Thurston Herricks, Evangelos Simeonidis, Nitin S. Baliga, John D. Aitchison, and Nathan D. Price. Combining inferred regulatory and reconstructed metabolic networks enhances phenotype prediction in yeast. *PLOS Computational Biology*, 13(5):e1005489, may 2017.

8 Danksagung

Herzlich möchte ich mich zunächst bei Dr. Friedemann Uschner bedanken, der mich bei meiner Bachelor-Arbeit betreut hat. Ohne ihn hätte ich mehrere technische Probleme, die der Modellierungsprozess mit sich bringt, nicht lösen können. Des Weiteren möchte ich mich bei Prof. Dr. Dr. h.c. Edda Klipp für die Bereitstellung der Originaldatei des Ionenmodels in Copasi bedanken, worüber ich einen Fehler im entsprechenden Paper auffinden konnte, was zur erfolgreichen Implementierung des Modells schließlich führte. Jorin Diemer danke ich dafür, dass er mich auf einen Programmierfehler hinwies und er mir seine Gedanken und Schriften zum Ionenmodel darbot.
Bei Björn Goldenbogen möchte ich mich für die Bereitstellung des Volumenmodels bedanken, welches sich noch im Veröffentllichungsprozess befindet. Mein allgemeiner Dank gilt der Gruppe von Edda Klipp, die mich freundlich in ihren Räumlichkeiten empfangen haben und mit vielen kleinen Unterstützungen mir die Bachelor-Arbeit erleichtert haben. Danken möchte ich meinen Eltern, die mich Zeit meines Lebens, auch wenn es mal ungewöhnlich schwer wurde, mit all ihrer Kraft liebevoll unterstützt haben.
Besonders möchte ich mich abschließend bei der Liebe meines Lebens, meiner Verlobten, bedanken. Du musstest mit mir jetzt schon ungeheure Tiefen durchgehen, die bei mir in den letzten Jahren aufgeklafft sind. Wir beide wissen, dass diese sich wohl nie mehr schließen werden können. Dank dir habe ich aber einen unheimlich starken Willen zum Kämpfen entwickelt, womit ich jede Dunkelheit durchstehen kann.