

Interaction between Experiment, Modeling and Simulation of Spatial Aspects in the JAK2/STAT5 Signaling Pathway

E. Friedmann, A. C. Pfeifer, R. Neumann, U. Klingmüller and R. Rannacher

Abstract Fundamental progress in systems biology can only be achieved if experimentalists and theoreticians closely collaborate. Mathematical models cannot be formulated precisely without deep knowledge of the experiments while complex biological systems can often not be understood fully without mathematical interpretation of the dynamic processes involved.

In this article, we describe how these two approaches can be combined to gain new insights on one of the most extensively studied signal transduction pathways, the Janus kinase (JAK)/ signal transducer and activator of transcription (STAT) pathway. We focus on the parameters of a model describing how STAT proteins are transported from the membrane to the nucleus where STATs regulate gene expression. We discuss which parameters can be measured experimentally in different cell types and how the unknown parameters are estimated, what the limits of these techniques and how accurate the determinations are.

Keywords: JAK2/STAT5 Signaling Pathway, Reaction Diffusion Equations, Parameter Estimation

MSC 2000

E. Friedmann, R. Neumann, R. Rannacher
Department of Applied Mathematics, University Heidelberg, Im Neuenheimer Feld 293, 69120 Heidelberg, Germany, e-mail: friedmann@iwr.uni-heidelberg.de, e-mail: rebecca.neumann@iwr.uni-heidelberg.de

A. C. Pfeifer, U. Klingmüller
Division Systems Biology of Signal Transduction, DKFZ-ZMBH Alliance, German Cancer Research Center, Im Neuenheimer Feld 280, 69120 Heidelberg, Germany e-mail: a.pfeifer@dkfz-heidelberg.de

1 Biological Question

In multicellular organisms communication between cells is frequently mediated by signal molecules secreted to the extracellular space which bind to cell surface receptors. The signal has to be transmitted from the extracellular domain of the cell surface receptors to the nucleus and thereby regulates gene expression.

One of the most extensively studied signal transduction pathways is the JAK/STAT pathway [5]. Several members of the signal transducer and activator of transcription (STAT) protein family have been implicated in various cancers. Briefly, after binding of ligand to the receptor two receptor associated Janus kinases (JAK) transphosphorylate each other and subsequently tyrosine phosphorylate the cytoplasmic domain of the receptor. STAT proteins can then bind to the phospho-tyrosine residues via their SH2 domains and are phosphorylated by JAK. Phosphorylated STATs dissociate from the receptor, dimerize, move to the nucleus and regulate transcription of target genes.

How transport from the site of STAT phosphorylation at the plasma membrane to the site of action in the nucleus is mediated is still unclear. Whether STATs freely diffuse through the cytoplasm to reach the nuclear envelope or are actively transported along the cytoskeleton remains a matter of debate. Moreover, it is not known if STATs can in addition be phosphorylated by membrane-bound kinases on endosomes present in the cytosol which would reduce the distance between the site of phosphorylation and nuclear envelope.

Here, we investigated the erythropoietin (Epo)-regulated JAK2/STAT5 pathway dynamics in two geometrically different cell types to address these questions. The fibroblast cell line NIH3T3 grows attached to surfaces, forming several tens of microns long and often branched stretches of cytoplasm surrounding an elliptical nucleus of approximately 10 μm in diameter, resulting in potentially large distances between plasma membrane and nuclear envelope. On the other hand, CFU-E (colony-forming unit erythroid stage) cells are primary cells isolated from mouse embryonic livers. They are precursors of red blood cells, growing in suspension not attached to surfaces, and showing a spherical shape with a diameter of roughly 10 μm for the whole cell and 8 μm for the nucleus, resulting in a fairly small cytoplasmic volume and only a short distance from plasma membrane to the nuclear envelope (Fig. 1).

2 Biological Data

A major limitation in systems biology remains the lack of sufficient high-quality, quantitative data for different variables of systems under investigation. To overcome this constraint, we have based our mathematical modeling on experimental data acquired by different experimental techniques, all generating quantitative data of high quality. Nevertheless, the restrictions of each method have to be assessed carefully to avoid misinterpretations of data. In addition, it is advisable to establish

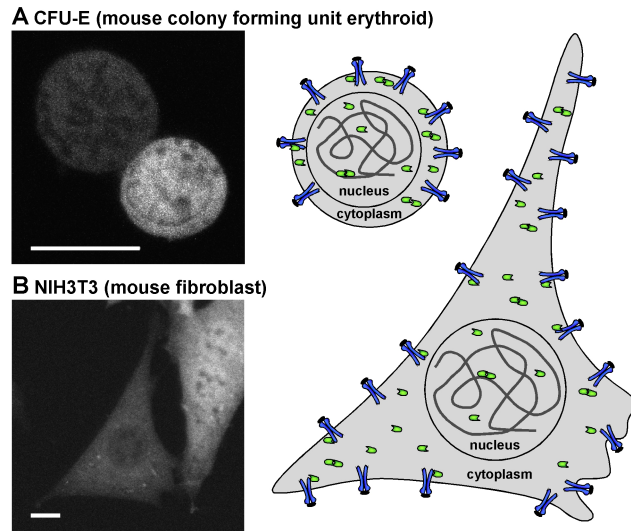


Fig. 1: Two different cell types used for this study. (A) CFU-E cells and (B) NIH3T3-EpoR cell expressing STAT5-GFP imaged on a confocal microscope. Scale bar, 10 μm . Schematic representations of the cells also indicate localization of EpoR (blue) on the plasma membrane with Epo (black) bound as well as STAT5 (green) in the cytoplasm in monomeric and dimeric form as well as bound to the receptor. Nuclear STAT5 is also shown bound to chromatin (dark grey line). JAK2 is omitted for simplification. Cells are drawn approximately to scale, proteins are overrepresented.

standard procedures for cell culture, sample preparation and experimental setup to guarantee comparable results ([9]).

2.1 Quantitative Immunoblotting Data

To measure dynamic changes in total protein concentrations as well as transient protein modifications such as tyrosine phosphorylation immunoblotting is commonly used. Proteins in cell extracts are separated by gel electrophoresis. Subsequently, they are electrophoretically transferred to a membrane on which they can be detected with antibodies specific for the protein or protein modification. Under certain circumstances it is necessary to purify the protein of interest from the remaining cell extract prior to immunoblotting by immunoprecipitation. This is especially useful, if proteins are only present in the cell at low concentrations or less specific antibodies are used for immunoblotting. For EpoR, JAK2 and STAT5 a combination of immunoprecipitation and immunoblotting was used. To reduce the error of the procedure

recombinant proteins were used as calibrators ([7]). This allows to determine a rough estimate of the number of protein molecules per cell.

When necessary, a crude separation of cytoplasm and nucleoplasm can be achieved by taking advantage of the different detergent sensitivities of the outer cell membrane and the nuclear envelope. However, this biochemical separation is never complete and leads to losses of certain fractions of the compartments that are difficult to determine. Furthermore, large numbers of cells are required if quantitative immunoblotting and immunoprecipitation are combined (on average one million cells per data point). If additionally a separation of cellular compartments, e.g. the cytoplasm and the nucleus, is required the cell number has to be further increased to obtain data of sufficient quality. In the case of established cell lines this usually only poses a handling problem, but if primary cells like CFU-E cells are used, experimentalists are strongly limited by the number of animals that can and should be sacrificed for an experiment. Furthermore, these biochemical methods only yield average data from large cell population. Asynchronous dynamics of single cells can be lost through averaging.

2.2 *Live Cell Imaging*

For the presented study it was necessary to determine the cell size and shape of CFU-E and NIH3T3 cells in addition to the biochemically measured dynamic changes of protein concentrations and protein modifications. Cellular features such as cell size and shape or more detailed information on protein localization can only be assessed by microscopy which offers a high spatial resolution at the single cell level.

We have determined the size of CFU-E cells and the CFU-E cell nucleus by transmitted light microscopy, measuring the diameter of the cell and the nucleus and assuming the cell shape to be a perfect sphere. For NIH3T3 cells, cells were enzymatically detached from their growth surfaces and assumed to adopt spherical shape in suspension and also measured by transmitted light microscopy. Alternatively, NIH3T3 cells expressing fluorescently labeled protein were imaged by confocal microscopy. For volume measurements by confocal microscopy a series of images were acquired, moving the sample in z-direction. This image stack covering the whole cell in x, y and z was analyzed and the volume calculated from the number of voxels. Thus, by expressing fluorescently labeled proteins with specific localizations the sizes of different cellular compartments can be determined.

To measure the changes of STAT5 localization over time after addition of Epo, STAT5 labeled with green fluorescent protein (STAT5-GFP) was followed in single cells by timelapse microscopy (Fig. 8). This allows to quantify the ratio of nuclear to cytosolic protein, detect potential concentration inhomogeneities as well as absolute protein concentrations if properly calibrated. The sampling rate that can be achieved is more than one image per minute if necessary. Sampling rate as well as duration of the experiment have to be adjusted depending on the goal of the experiment.

In general, again depending on the strength of the signal, 200-500 images can be acquired of a sample, with a sampling rate between less than a second up to hours.

However, a few restrictions have to be taken into account. Live cell imaging requires the expression of fluorescently labeled proteins. This is usually achieved by genetically tagging the protein of interest with a fluorescent protein and expressing it in addition to the unlabeled protein already present in the cells. This can lead to severe alterations in the pathway dynamics and has to be tested carefully. For microscopy, cells have to be immobilized on thin optical glass surfaces. This is easily done with adherent cells like the NIH3T3 cells by growing them in culturing dishes with a glass bottom and directly placing the culturing dish on the microscope, usually equipped with an incubation chamber to maintain the appropriate environment for the cells. In the case of cells growing in suspension such as the CFU-Es imaging of living cells over time is rather challenging and was not done for this study. In the future, this might be possible as commercial solutions for imaging suspension cells are being developed.

2.3 Measurement of Diffusion and Transport Kinetics

2.3.1 Fluorescence Correlation Spectroscopy (FCS)

The protein localization in cells in steady state conveys a very static picture. However, many proteins are constantly in rapid movement. Fluorescence Correlation Spectroscopy (FCS) is a widely applied method to measure movement of proteins caused by diffusion or reaction-diffusion systems. Fluctuations of fluorescently-labeled proteins within the focal volume of a confocal or two-photon microscope are recorded and analyzed by temporal autocorrelation.

To achieve a high level of fluctuations, the concentration of the observed protein should be kept low. This is often not the case for standard mammalian expression systems but can be accomplished by using an inducible expression system. Alternatively, a large fraction of the fluorescence can be photobleached before the FCS measurement. However, this can cause artefacts due to photodamage caused by the bleaching process. Furthermore, spatial inhomogeneities of the cell as well as photophysical characteristics of the fluorophore can cause difficulties during data acquisition and analysis. This can lead to a fairly high variability of parameters determined from FCS data.

The diffusion coefficient of STAT5-GFP in the cytoplasm of fibroblasts was measured by FCS and determined to be approximately $15 \mu\text{m}^2/\text{s}$. No significant difference for the diffusion coefficient of STAT5-GFP in starved cells and cells stimulated by Epo could be detected. For CFU-E cells a similar diffusion coefficient was assumed.

2.3.2 Fluorescence Recovery After Photobleaching (FRAP)

In the case of FCS usually only faster processes such as diffusion or binding kinetics can be studied. In addition, proteins can also be transported from one cellular compartment to another. These protein dynamics in addition to more rapid molecule movements - can be assessed by fluorescence recovery after photobleaching (FRAP). In a typical FRAP experiment, the steady state fluorescence distribution is perturbed by photobleaching part of the fluorescence with a strong laser pulse and the exchange of bleached and unbleached proteins is monitored over time.

Here, we have measured nuclear import and export rates of unphosphorylated STAT5 in NIH3T3 cells ([5] and fig. 8). As standard FRAP analysis is only applicable to steady state situations at least on the time scale of the experiment and recovery exchange between nucleus and cytoplasm is on a similar time scale as protein phosphorylation after Epo stimulation only nucleocytoplasmic shuttling of unphosphorylated STAT5 in serum starved cells was measured. FRAP was not performed in CFU-E cells as this would require stable maintenance of non-adherent cells on the microscope to allow for analysis of the measurement as discussed above.

3 Modeling

In this section we will consider two models describing the same biological processes, one without spatial resolution described by a system of ordinary differential equations (ODE), and the other with partially considered spatial resolution which is described by a mixed system containing ordinary and partial differential equations (ODE and PDE). Parameter estimation was performed for the ODE model and the same parameter values are used in the second model.

3.1 Ordinary Differential Equations Model

In our models, we focus only on a part of the JAK2/STAT5 signaling pathway: after binding of the hormone erythropoietin (Epo) to its receptor (EpoR), STAT5 is phosphorylated at the EpoR by JAK2 with rate r_{act} , dimerizes and diffuses through the cytoplasm to the nucleus. STAT5 is imported into the nucleus with rate r_{imp} and pSTAT5 with rate r_{imp2} . The processes in the nucleus are described by four linear delay equations including the dephosphorylation of phosphorylated STAT5 which then is exported back to the cytoplasm with rate r_{exp} .

We denote by u_0 the concentration of unphosphorylated STAT5 in the cytoplasm, by u_1 the concentration of the phosphorylated STAT5 in the cytoplasm, by u_2 the concentration of the unphosphorylated STAT5 in the nucleus and by u_3 the concentration of the phosphorylated STAT5 in the nucleus. The variables u_4, \dots, u_7 are introduced

to describe the processes in the nucleus by linear delay equations, they are so called fictitious concentrations.

$$u'_0(t) = -\frac{r_{act}}{v_{cyt}} \cdot pJAK(t) \cdot u_0(t) - \frac{r_{imp}}{v_{cyt}} \cdot u_0(t) + \frac{r_{exp}}{v_{cyt}} \cdot u_2(t) \quad (1)$$

$$u'_1(t) = \frac{r_{act}}{v_{cyt}} pJAK(t) \cdot u_0(t) - \frac{r_{imp2}}{v_{cyt}} \cdot u_1(t) \quad (2)$$

$$u'_2(t) = \frac{r_{delay}}{v_{nuc}} \cdot u_7(t) - \frac{r_{exp}}{v_{nuc}} \cdot u_2(t) + \frac{r_{imp}}{v_{nuc}} \cdot u_0(t) \quad (3)$$

$$u'_3(t) = \frac{r_{imp2}}{v_{nuc}} \cdot u_1(t) - \frac{r_{delay}}{v_{nuc}} \cdot u_3(t) \quad (4)$$

$$u'_4(t) = \frac{r_{delay}}{v_{nuc}} (u_3(t) - u_4(t)) \quad (5)$$

$$u'_5(t) = \frac{r_{delay}}{v_{nuc}} (u_4(t) - u_5(t)) \quad (6)$$

$$u'_6(t) = \frac{r_{delay}}{v_{nuc}} (u_5(t) - u_6(t)) \quad (7)$$

$$u'_7(t) = \frac{r_{delay}}{v_{nuc}} (u_6(t) - u_7(t)) \quad (8)$$

This model was considered for two different cell types: a spherical shaped CFU-E and a NIH3T3 fibroblast cell. Therefore we obtain two sets of parameters, one for each cell type: The initial values are

$$u_1(0) = u_3(0) = u_4(0) = u_5(0) = u_6(0) = u_7(0) = 0 \quad (9)$$

$$\text{CFU-E: } u_0(0) = 50 \hat{a} \cdot \text{mol}/\mu\text{m}^3 \quad (10)$$

$$u_2(0) = 18 \hat{a} \cdot \text{mol}/\mu\text{m}^3 \quad (11)$$

$$\text{NIH3T3: } u_0(0) = 16 \hat{a} \cdot \text{mol}/\mu\text{m}^3 \quad (12)$$

$$u_2(0) = 20 \hat{a} \cdot \text{mol}/\mu\text{m}^3, \quad (13)$$

whereas the number of molecules per compartment was determined using a combination of immunoprecipitation and immunoblotting as described in paragraph 2.1. Here, \hat{a} is the Avogadro constant, so that the unit $\hat{a} \cdot \text{mol}$ is the number of molecules.

The parameters v_{cyt} and v_{nuc} which represent the average volume of the cytoplasm and nucleus were measured by transmitted light microscopy (section 2.2). We use the values: $v_{cyt} = 429 \mu\text{m}^3$, $v_{nuc} = 268 \mu\text{m}^3$ for the CFU-E cell and $v_{cyt} = 1758 \mu\text{m}^3$, $v_{nuc} = 366 \mu\text{m}^3$ for the NIH3T3 cell. The nuclear import and export rates can be measured only for the unphosphorylated STAT5 in NIH3T3 cells by FRAP experiments (section 2.3.2): $r_{imp} \approx 0,1 \text{ s}^{-1}$ and $r_{exp} \approx 0,09598 \text{ s}^{-1}$. The import rate of the unphosphorylated STAT5 in the CFU-E cells is assumed to be approximately the same as in the NIH3T3 cells, so that the same value for r_{imp} ($\approx 0,1 \text{ s}^{-1}$) is used for both cell types in our model.

Another input function in this model is the phosphorylation function $pJAK(t)$. In the experiments the STAT5 molecules are phosphorylated through a controlled

input of Epo which activates the receptor-associated kinase JAK2 which then phosphorylates STAT5. The biological processes at the receptor can be modeled as an additional receptor module which makes our systems of equations and the parameter set much larger. Since the concentration of phosphorylated JAK2 molecules can be measured we simplify our model and use the data points for pJAK2 interpolated with a smooth spline as input (figure 3). For our model we assume that the number of JAK2 molecules at the plasma membrane is equal to the number of EpoR molecules on the plasma membrane.

Our system of equations (1)-(13) is a homogenous linear system which can be formulated in a general form as follows:

$$u'(t) = f(t, u(t)) \quad (14)$$

$$u(0) = u_0. \quad (15)$$

Theorem 1. *The system (14)-(15) is well-posed and numerically stable.*

The proof of this theorem is based on standard techniques. The right hand side $f(t, u(t))$ is Lipschitz, so that we can apply Picard-Lindelöf's theorem and get existence and uniqueness locally. The local result can then be extended to a globally defined solution. The function $f(t, u(t))$ fulfills an additional monotonicity condition which gives the stability. The simulations are then performed by applying the implicit Euler or Crank-Nicolson method.

3.2 Parameter estimation

Parameters that can not be experimentally measured for our model for the CFU-E cell are the phosphorylation rate of STAT5, its export rate, the import rate of phosphorylated STAT5 and the time delay for the processes in the nucleus: r_{act} , r_{exp} , r_{imp2} and r_{delay} . For the model of a NIH3T3 cell we have only three unknown parameters: r_{act} , r_{imp2} and r_{delay} . To determine these unknown parameters, parameter estimation was performed using the software PottersWheel [3], developed to perform data-based modeling of partially observed and noisy systems like signal transduction pathways.

For each cell type two series of experiments were performed. For the CFU-E cells the observables in both experiments were $pJAK(t)$, $u_1(t)$, $u_0(t) + u_1(t)$, $u_2(t) + u_3(t)$. In the experiments with the NIH3T3 cells either $pJAK(t)$, $u_1(t)$, $u_3(t)$ and $u_2(t) + u_3(t)$ were observed, or $pJAK(t)$, $u_0(t)$, $u_1(t)$, $u_2(t)$ and $u_3(t)$. The data for the parameter estimation were generated by quantitative immunoblotting where the specific components could be identified via antibodies. For an interpretation of the errors of biochemical data see [7].

The parameter estimation process belongs to the group of non-linear least square problems. The merit function χ^2 which is optimized in PottersWheel to fit the model $y = y(t; p)$ is given by

$$\chi^2(p) = \sum_{i=1}^N \left(\frac{y_i - y(t_i; p)}{\sigma_i} \right)^2, \quad (16)$$

with y_i being data point i with standard deviation σ_i and $y(t_i; p)$ being the model value at time point i for parameter values p . As the measurement errors are normally distributed, the minimization of the weighted least-square error corresponds to applying a Maximum Likelihood estimator for the unknown parameters [3]. For the fitting one can choose between powerful deterministic and stochastic optimization algorithms. Our data were fitted 500 times in a fit sequence using the same dataset where the initial value in the consequent fit is chosen as the best fit in the previous sequence with parameter disturbed by a random number. Only the best 30% percent have been analyzed further. The results are listed in table 1 and 2 (first column).

For the interpretation of the simulation results it is important to determine the relative deviation of the parameters: For the NIH3T3 fibroblast cell the relative deviation is less than 10% for the parameter describing the phosphorylation of STAT5 and up to 20 % for the delay time in the delay reactions considered. For the parameter set of the CFU-E cell the smallest relative deviation (0.61%) is obtained for the export rate of unphosphorylated STAT5 and the largest relative deviation is obtained for the phosphorylation rate (27%). The relative deviation of the time delay remains approximately the same for both models. How these fits could be further improved is discussed in section 5.

	CFU-E	CFU-E min	CFU-E max
$r_{\text{act}} (\text{min}^{-1})$	11	9	15
$r_{\text{imp2}} (\text{min}^{-1})$	58	47	71
$r_{\text{exp}} (\text{min}^{-1})$	225	194	260
$r_{\text{delay}} (\text{min}^{-1})$	265	263	266

Table 1: Deviation rates of the parameters used in the model for the CFU-E cell.

	NIH3T3	NIH3T3 min	NIH3T3 max
$r_{\text{act}} (\text{min}^{-1})$	187	171	204
$r_{\text{imp2}} (\text{min}^{-1})$	1010	886	1151
$r_{\text{delay}} (\text{min}^{-1})$	194	162	232

Table 2: Deviation rates of the parameters used in the model for the NIH3T3 cell.

The graph in figure 3 represents an exemplary dataset from NIH3T3 cells, with the result of the parameter estimation shown in blue. Multi-experiment fitting was performed with PottersWheel.

3.3 Reaction Diffusion Model

To address the biological question described above we add diffusion for unphosphorylated and phosphorylated STAT5 in the cytoplasm. Different transport processes can be modeled by different diffusion coefficients. At first, we model free diffusion using a constant diffusion coefficient based on measurements by fluorescence correlation spectroscopy (FCS, section 2.3.1). The diffusion coefficient $D = 15 \mu\text{m}^2/\text{s} = 900 \mu\text{m}^2/\text{min}$ was used in the simulations. The additional transport of the molecules along the microtubules is modeled for the NIH3T3 cell through an anisotropic diffusion coefficient whereas the mainstream direction of STAT5 movement was set in the y-direction of the cell.

To model the reaction diffusion equations we consider two different geometries, spherical for a CFU-E cell and a long tube with an ellipsoidal body for a NIH3T3 cell (Fig. 4). To answer the biological question only the cytoplasm has to be dissolved spatially. The processes in the nucleus such as DNA binding and dephosphorylation of STAT5 do not have to be known in detail. For their description it is sufficient to use time delays as black box elements. As already described in the previous section, phosphorylation as well as nuclear import and export of STAT5 only occurs on the boundary of our considered computational domain, the cytoplasm. For this specific question we therefore obtain a mixed system of differential equations: two diffusion equations with Robin boundary conditions and six ODE, two of them are coupled to the PDEs through the import terms and the other four describe the processes in the nucleus by linear delay equations:

Cytoplasm: Ω_{cyt}

$$\partial_t u_0(t, x) = D \Delta u_0(t, x) \quad (17)$$

$$\partial_t u_1(t, x) = D \Delta u_1(t, x) \quad (18)$$

Nucleus: Ω_{nuc}

$$u_2'(t) = \frac{r_{\text{delay}}}{v_{\text{nuc}}} \cdot u_7(t) - \frac{r_{\text{exp}}}{v_{\text{nuc}}} \cdot u_2(t) + \frac{r_{\text{imp}}}{v_{\text{nuc}}} \cdot \frac{1}{|\partial \Omega_{\text{nuc}}|} \cdot \int_{\partial \Omega_{\text{nuc}}} u_0(t, s) \, ds \quad (19)$$

$$u_3'(t) = \frac{r_{\text{imp2}}}{v_{\text{nuc}}} \cdot \frac{1}{|\partial \Omega_{\text{nuc}}|} \cdot \int_{\partial \Omega_{\text{nuc}}} u_1(t, s) \, ds - \frac{r_{\text{delay}}}{v_{\text{nuc}}} \cdot u_3(t) \quad (20)$$

$$u_4'(t) = \frac{r_{\text{delay}}}{v_{\text{nuc}}} (u_3(t) - u_4(t)) \quad (21)$$

$$u_5'(t) = \frac{r_{\text{delay}}}{v_{\text{nuc}}} (u_4(t) - u_5(t)) \quad (22)$$

$$u_6'(t) = \frac{r_{\text{delay}}}{v_{\text{nuc}}} (u_5(t) - u_6(t)) \quad (23)$$

$$u_7'(t) = \frac{r_{\text{delay}}}{v_{\text{nuc}}} (u_6(t) - u_7(t)) \quad (24)$$

The initial conditions are the same used in the previous section, we only have to pay attention that the concentrations of unphosphorylated and phosphorylated STAT molecules in the cytoplasm are now space dependent ($u_0(t, x)$ and $u_1(t, x)$). The phosphorylation, import and export of molecules enter through the Robin boundary conditions:

$$D\partial_n u_0(t, \partial\Omega_{\text{cyt}}) = -\frac{r_{\text{act}}}{|\partial\Omega_{\text{cyt}}|} \cdot \text{pJAK}(t, x) \cdot u_0(t, x) \quad (25)$$

$$D\partial_n u_0(t, \partial\Omega_{\text{nuc}}) = -\frac{r_{\text{imp}}}{|\partial\Omega_{\text{nuc}}|} \cdot u_0(t, x) + \frac{r_{\text{exp}}}{|\partial\Omega_{\text{nuc}}|} \cdot u_2(t) \quad (26)$$

$$D\partial_n u_1(t, \partial\Omega_{\text{cyt}}) = \frac{r_{\text{act}}}{|\partial\Omega_{\text{cyt}}|} \cdot \text{pJAK}(t, x) \cdot u_0(t, x) \quad (27)$$

$$D\partial_n u_1(t, \partial\Omega_{\text{nuc}}) = -\frac{r_{\text{imp2}}}{|\partial\Omega_{\text{nuc}}|} \cdot u_1(t, x), \quad (28)$$

whereas $\partial\Omega_{\text{cyt}}$ represents the outer boundary of the cell, i.e. the membrane and $\partial\Omega_{\text{nuc}}$ the boundary of the nucleus.

At this point it is important to note that the diffusion coefficient is contained in the measured parameters and that the input curve for the phosphorylation of the STAT5 molecules ($\text{pJAK}(t)$) must be recalculated in a form of concentration using the number of receptors of the specific cell type. For the CFU-E cell we have:

$$\text{pJAK}(t, x) = \begin{cases} 1.244 \cdot \text{pJAK}(t) & x \in \partial\Omega_{\text{cyt}} \\ 0 & \text{elsewhere} \end{cases}$$

and for the NIH3T3 cell:

$$\text{pJAK}(t, x) = \begin{cases} 5.65 \cdot \text{pJAK}(t) & x \in \partial\Omega_{\text{cyt}} \\ 0 & \text{elsewhere.} \end{cases}$$

In this input the number of receptor molecules on the cell surface is hidden. For CFU-E cells the maximum number of ligand binding sites, i.e. available receptor dimers at the cell surface, has been reported to be approximately 1000 ([11]) yielding 2000 receptor molecules at the surface of a CFU-E cell. In NIH3T3 cells we overexpress the receptor, presumably generating a higher density of receptor at the cell surface. As the number of cell surface receptors is difficult to measure we have assumed that the receptor density is similar to that in another cell line with overexpressed receptor (BaF3-EpoR, V. Becker, personal communication) and have assumed 28000 receptor molecules at the surface of a NIH3T3 cell for our model.

For the Reaction-Diffusion-System (14)-(25) we have to introduce the corresponding functional spaces which will assure the solvability: Let X denote a real Banach space with Norm $\|\cdot\|$. Ω is an open and bounded subset in \mathbb{R}^n .

Definition 1. The space $L^p((0, T), X)$ consists of all strongly measurable functions $u : [0, T] \rightarrow X$ with

$$\|u\|_{L^p((0, T), X)} := \left(\int_0^T \|u(t)\|^p dt \right)^{\frac{1}{p}} < \infty$$

for $1 \leq p < \infty$, and

$$\|u\|_{L^\infty((0, T), X)} := \text{esssup} \|u(t)\| < \infty.$$

For our system $X = H_0^1(\Omega)$, H^{-1} is the dual space of X and we have

$$H_0^1(\Omega) \subset L^2(\Omega) \subset H^{-1}(\Omega).$$

The well-posedness of the system is shown by following theorem:

Theorem 2. *For constant, positive and continuous parameters, the Reaction-Diffusion System from above, written in general form as*

$$\begin{aligned} \partial_t u(t, x) &= D\Delta u(t, x) + f(t, u) \\ D\partial_n u(t, s) + au(t, s)|_{\partial\Omega} &= g \\ u(0) &= u_0 \end{aligned}$$

has a unique solution $u_0, u_1 \in L^2((0, T), H^1(\Omega))$, $\partial_t u_0, \partial_t u_1 \in L^2((0, T), (H^1(\Omega))^{-1})$ and $u_2, \dots, u_7 \in C^1[0, T]$, where $u = (u_0, u_1, \dots, u_7)$.

The proof of this theorem can be done in a standard way by decoupling the equations for u_i , $i = 0, 1, 3, \dots, 7$ and applying the Banach's Fixpoint Theorem to u_2 . It can be found in [4] and in [1].

The numerical simulations are realized in our group intern software Gascoigne based on Finite Elements and contain grid generation, discretization in bilinear Finite Elements (Q_1) and the solution via the Rothe-Method and Crank-Nicolson, whereas the resulting system is solved by an multigrid algorithm.

4 Results

The results of the simulations of (1)-(13), dotted line, and (14)-(25), solid line, are plotted for CFU-E and NIH3T3 in Figure 5 and 6, respectively. There, we visualized the differences between the time distribution of the concentration of unphosphorylated STAT5 and phosphorylated STAT5 from the system without diffusion with $u_{i,h}(t) = \int_{\Omega_{\text{cyt}}} u_i(t, x) dx$, $i = 0, 1$, where $u_0(t, x), u_1(t, x)$ are solutions of the system with diffusion. In CFU-E cells, no effect of diffusion on the time distribution of phosphorylated STAT5 concentration was observed. There, the distance from the cell membrane to the nucleus is very short (Fig. 1), therefore, addition of diffusion of cytoplasmic STAT5 does not alter the results. The model (1)-(13) describes

the considered pathway in the CFU-E cells very well (Fig. 5). In NIH3T3 cells, adding diffusion of cytoplasmic STAT5 caused a higher time distribution of the concentration of phosphorylated STAT5 in the cytoplasm ($u_1(t)$) as well as a different steady state (Fig. 6). This is due to the long distance which the molecules have to cover to get from the membrane to the nucleus. The higher concentration of phosphorylated STAT5 in the cytoplasm is compensated with a lower concentration of phosphorylated STAT5 in the nucleus which indicates that the addition of diffusion does not change the amount of phosphorylated STAT5 in the whole cell.

For NIH3T3 cells the discrepancy between the model with diffusion and without diffusion is low (smaller than 1 molecule/ μm^3 , figure 6). To answer the biological question we simulated various transport processes considering the same parameters in each model (data not shown). Each transport process shows a distinct concentration development during time. Only because of fast diffusion the differences are very small. Also, the cell geometry influences the concentration development of phosphorylated STAT5 molecules during time. Here, the differences are more visible, even with fast diffusion. We will address the exact analyzation of these simulations in a future paper.

5 Perspectives

An exact interpretation of the simulation results can be obtained only after analyzing the reliability of the parameters. Sensitivity analysis is used to determine how sensitive a model is to changes in the value of the parameters of the model and to changes in the structure of the model. We perform a series of tests in which we set different parameter values to see how a change in the parameter causes a change in the dynamic behavior of our model. The model behaviour responds to changes in the parameters is shown in figure 7. There, the absolute deviations between the models is presented. Calculation of the relative deviation let us conclude that the effect of the relative deviation of the parameters on the concentration development is of the same order or even greater than the effect of the diffusion. The level of accuracy of our parameters is not sufficient. Thus the question which transport process is used by the STAT5 molecules to shuttle to the nucleus remains open.

Additionally, the limitations of the experimental techniques make the interpretation of the results difficult. Simulation results and experimental results are not always directly comparable as illustrated by the biochemical experiments. The simulations result in concentrations of the involved pathway components in a single cell whereas the biochemical data is generated from a large cell population and is affected by losses of cellular components due to the biochemical purification. It yields only averaged, relative concentrations that are difficult to convert to absolute numbers. A direct comparison of the observables from experiments with the ones from simulations was therefore not possible.

The detailed study of the two models give us some indications under which conditions an answer of the biological question might be possible. For similar linear

models in signal transduction we can conclude which geometry and which diffusion coefficient is necessary so that diffusion plays an important role in the dynamics of the observed pathway. First steps towards experimental design were established here: For the considered pathway we have to reduce the relative deviation of the parameters used in the model in order to better answer the biological question. One possibility is to add the observable $\frac{u_2(t)+u_3(t)}{u_0(t)+u_1(t)}$ from live cell imaging experiments (figure 8) to the estimation of the parameters. In microscopy, a better time resolution can be easily achieved which should have a positive effect on the relative deviation of the parameters. Recently, spatially resolved diffusion times for fluorescently labeled molecules in cells were measured by diffusion imaging microscopy ([6]). Similar measurements to investigate possible anisotropic diffusion of STAT5 could be performed. Experimental design could be further optimized to determine which measurements should be done in order to obtain better results. Furthermore, new biological questions could arise from the observed model behavior like: is STAT5 phosphorylated on the appendices of fibroblasts, are there active Epo receptors at all?

In this paper our aim was to emphasize the importance of interdisciplinary research on the example of the JAK2/STAT5 signaling pathway. High-quality quantitative measurements together with mathematical modeling, computational simulations, parameter estimation and experimental design helps to decide whether and which scientific question can be answered based on existing laboratory constraints.

6 Experimental Procedures

6.1 Cell lines and Preparation of primary CFU-E cells

All NIH3T3 cells were cultured in DMEM (Invitrogen) including 10% newborn calf serum. All media were supplemented with penicillin (100 U/ml) and streptomycin (100 mg/ml). CFU-E cells of d13.5 embryos from wild-type Balb/c mice were enriched as described [2].

NIH3T3-EpoR and NIH3T3-EpoR cells expressing STAT5-GFP under an inducible promoter were prepared as previously described [5].

Serum-starvation and stimulation of NIH3T3-EpoR cells for all experiments was performed by starving cells in DMEM supplemented with 25mM Hepes pH 7.4 (Invitrogen) for at least 5 hours and stimulating with 25 units/ml Epo (Janssen-Cilag Bad Homburg, Germany) at 37 °C for indicated times.

6.2 Quantitative Immunoblotting

For measuring activated JAK2 and STAT5 in NIH3T-EpoR cells, the cytoplasm of 3×10^6 cells per time point was lysed with 0.5% NP-40 lysis buffer supplemented with

aprotinin and AEBSF (Sigma-Aldrich). After removal of the cytoplasmic fraction the nuclei were lysed with 1% NP-40 lysis buffer. For immunoprecipitation, lysates were incubated with anti-EpoR antibodies (M-20, Santa Cruz, La Jolla, CA) and anti-JAK2 (Upstate) or anti-STAT5 antibodies (Santa Cruz). Immunoprecipitated proteins were loaded in randomised fashion on SDS polyacrylamide gel as described [8], separated by electrophoresis and immunoblotted using anti-phosphotyrosine monoclonal antibody 4G10 (Upstate Biotechnology, Lake Placid, NY) and secondary HRP coupled anti-mouse antibody (Amersham Biosciences, Freiburg, Germany). Typically, immunoblots were incubated with ECL substrate (Amersham Biosciences) for 1 minute and exposed for 5-10 minutes on a LumiImager (Roche Diagnostics, Mannheim, Germany). Alternatively, membranes were incubated with ECL advance (Amersham Biosciences) for 2 minutes and exposed for 30 seconds 2 minutes on a LumiImager. LumiAnalyst software (Roche Diagnostics) was used for quantification. Antibodies were removed by treating the blots with beta-mercaptoethanol and SDS as described. Reprobes were performed using anti-EpoR antibody (Santa Cruz) and anti-STAT5 antibody (Santa Cruz). Quantitative immunoblotting data was processed using GellInspector software. Normalisers SBP-JH2JH1 for pJAK2 and GST-STAT5b for pSTAT5 were used.

6.3 Live cell imaging, FRAP and FCS

Confocal imaging was performed on a Leica TCS SP5 equipped with an oil immersion objective $63 \times /1.4$ oil and the pinhole set to 1 Airy unit. FCS measurements were performed on a Leica TCS SP2 confocal microscope equipped with a water immersion objective $63 \times /1.2$ W CORR, the pinhole set to 1 Airy unit and an FCS unit (Leica Microsystems, Mannheim, Germany). For all time-lapse imaging and FCS measurements the sample and the microscope were held at 37°C by a climate chamber.

For live cell imaging STAT5-GFP expression in NIH3T3-EpoR-STAT5-GFP cells was induced with 10 ng/ml doxycycline 16 hours before a serum-starvation of 5 hours. After addition of 25 units/ml Epo cells were imaged on a confocal microscope for a minimum of 4 hours at 37°C . The ratio of nuclear to cytoplasmic STAT5-GFP concentration was determined by measuring the fluorescence intensity in the respective regions for each cell in ImageJ.

Nuclear import and export parameters of STAT5-GFP in NIH3T3-EpoR cells were determined by FRAP experiments as previously described [5].

For FCS measurements STAT5-GFP expression in NIH3T3-EpoR-STAT5-GFP cells was induced with 10 ng/ml doxycycline 16 hours before a serum-starvation of 5 hours. A minimum of two position per cell per compartment were measured by FCS in unstimulated cells or cells stimulated with 25 units/ml Epo for 15-45 minutes. Intensity fluctuations in the focal volume were analyzed by temporal autocorrelation. Diffusion coefficients were estimated by fitting a one-component model to the data.

References

1. Friedmann, E, Neumann, R. and Rannacher, R., *Well-posedness for a spatio-temporal model of the JAK2/STAT5 Signaling Pathway*, Diffusion and Association Processes in Biological Systems: Theory, Computation and Experiments, PMC Biophysics, submitted 2010.
2. Ketteler, R., Glaser, S., Sandra, O., Martens, U. M. and Klingmüller, U., *Enhanced transgene expression in primitive hematopoietic progenitor cells and embryonic stem cells efficiently transduced by optimized retroviral hybrid vectors*, Gene Therapy, 9, 477-487, 2002.
3. Maiwald, T. and Timmer, J., *Dynamical modeling and multi-experiment fitting with Potters-Wheel*, Bioinformatics 24(18):2037-43, 2008.
4. Neumann, R., *Räumliche Aspekte in der Signaltransduktion*, Diplomarbeit Ruprecht-Karls-Universität Heidelberg, 2009.
5. Pfeifer A.C., Kaschek D., Bachmann J., Klingmüller U. and Timmer J., *Model-based extension of high-throughput to high-content data*, BMC Syst Biol., 4:106 2010.
6. Roth C. M., Heinlein P. I., Heilemann M., and Hertel D.-P., *Imaging Diffusion in Living Cells Using Time-Correlated Single-Photon Counting*, Anal. Chem., 79 (19), pp 73407345, 2007.
7. Schilling M., Maiwald T., Bohl S., Kollmann M., Kreutz C., Timmer J. and Klingmüller U., *Computational processing and error reduction strategies for standardized quantitative data in biological networks*, FEBS J. 272(24):6400-11, 2005.
8. Schilling M., Maiwald T., Bohl S., Kollmann M., Kreutz C., Timmer J. and Klingmüller U., *Quantitative data generation for systems biology: the impact of randomisation, calibrators and normalisers*, Syst Biol 2005, 152(4):193-200, 2005.
9. Schilling M., Pfeifer A.C., Bohl S. and Klingmüller U., *Standardizing experimental protocols*, Curr Opin Biotechnol. 2008.
10. Schille P. and Hausteiner E., *Fluorescence Correlation Spectroscopy. An Introduction to its Concepts and Applications*, Annu Rev Biophys Biomol Struct. 36:151-69, 2007.
11. Yousoufian H., Longmore G., Neumann D., Yoshimura A. and Lodish H.F., *Structure, function, and activation of the erythropoietin receptor*, Blood. May 1;81(9):2223-36, 1993.
12. *High Performance Adaptive Finite Element Toolkit Gascoigne*, <http://gascoigne.uni-hd.de>.

Acknowledgements This work was supported by the Helmholtz Alliance on Systems Biology (SBCancer, Submodule V.7). ACP was supported by the German Federal Ministry of Education and Research (BMBF) grant FORSYS-ViroQuant (#0313923). We thank Julie Bachmann for providing biochemical data for the CFU-E cells and V. Becker for personal communication.

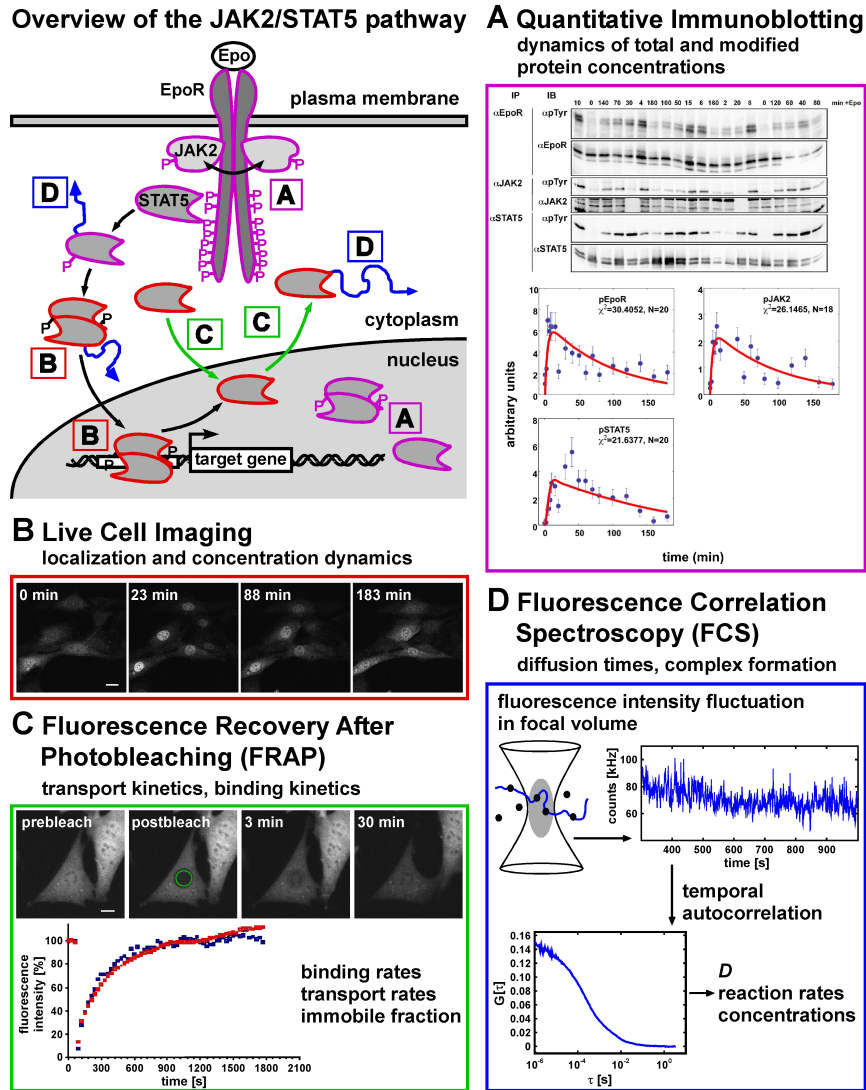


Fig. 2: Overview of experimental methods. Several quantitative techniques were combined to generate data for mathematical modeling. An overview of the JAK-STAT signaling pathway is shown in the top left panel. Colored frames and arrows represent the protein species and reactions monitored by the individual techniques. Bold letter in boxes refer to the respective techniques. (A) Cytoplasmic extracts of NIH3T3-EpoR cells were subjected to immunoprecipitation with antibodies against EpoR and JAK2, or STAT5, and analyzed by quantitative immunoblotting. Samples are randomized to reduce correlated errors. Processed data points (blue points) and fits of an ODE model to the data (red curve) are shown below. (B) Nuclear accumulation of STAT5-GFP in NIH3T3-EpoR cells after addition of Epo is investigated by timelapse microscopy on a confocal microscope. (C) Fluorescence recovery after photobleaching (FRAP) is used to analyze nucleocytoplasmic shuttling of STAT5-GFP. Nuclear STAT5-GFP in NIH3T3-EpoR cells was photobleached (green circle) and recovery of nuclear fluorescence was observed. Scale bar, 10 μm . Two representative processed data sets are shown for unphosphorylated STAT5-GFP. (D) Mobility of STAT5-GFP in the cytoplasm was measured by fluorescence correlation spectroscopy (FCS). Intensity fluctuations in the focal volume (represented in green) were analyzed by temporal autocorrelation. Diffusion coefficients were estimated by fitting a one-component model to the data.

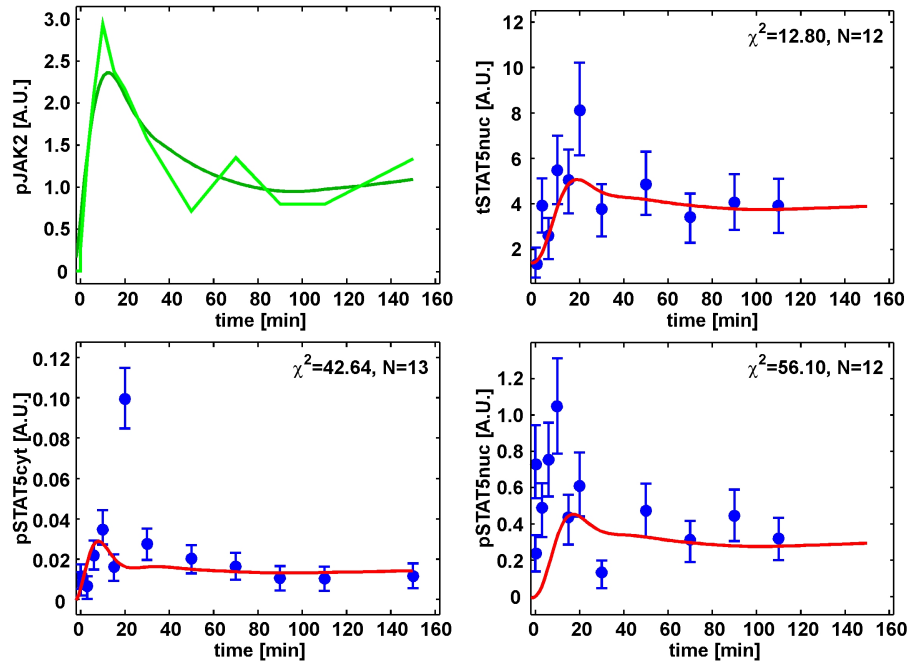


Fig. 3: Data-based ODE modeling of STAT5 phosphorylation in NIH3T3-EpoR cells. NIH3T3-EpoR cells were serum-starved for 5 hours and stimulated with Epo. Cells were fractionated in nuclear and cytoplasmic extracts. Extracts were subjected to immunoprecipitation with antibodies against EpoR and JAK2 (cytoplasm only), or STAT5 (cytoplasm and nucleus), and analyzed by quantitative immunoblotting. Processed data points (blue points) and fits of the mathematical model to the data (red curves) are shown above. A spline through the data points for phosphorylated JAK2 (upper left panel) was used as input.

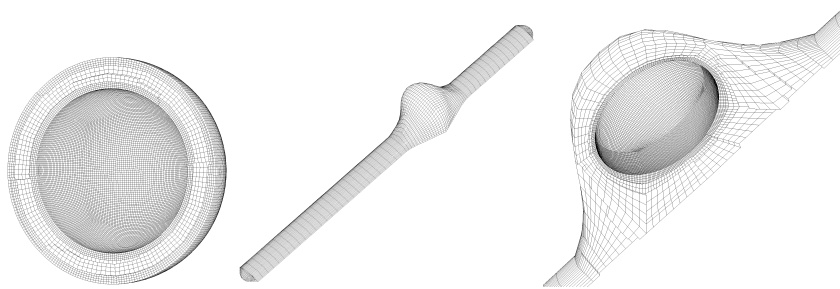


Fig. 4: Constructed grid for the CFU-E cell, NIH3T3 cell and a section through the nucleus of the NIH3T3 cell.

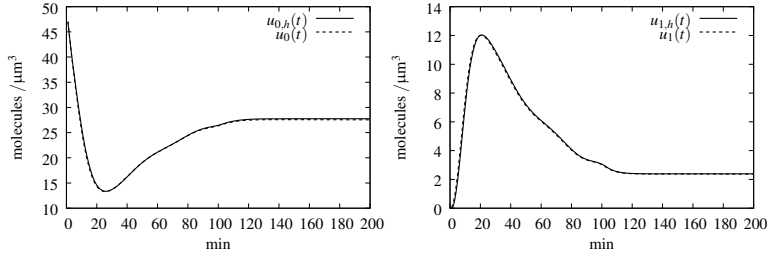


Fig. 5: The solution $u_0(t)$ (unphosphorylated STAT5) of system (1)-(13) in comparison to the solution $u_{0,h}(t) = \int_{\Omega_{\text{cyt}}} u_0(t,x)dx$ of system (14)-(25) (left) and the solution $u_1(t)$ (phosphorylated STAT5) of system (1)-(13) in comparison to the solution $u_{1,h}(t) = \int_{\Omega_{\text{cyt}}} u_1(t,x)dx$ of system (14)-(25) (right) in a CFU-E cell.

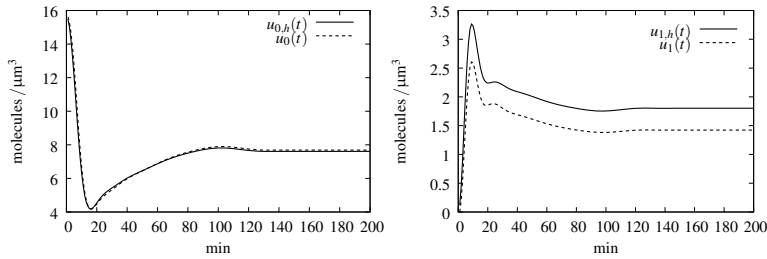


Fig. 6: The solution $u_0(t)$ (unphosphorylated STAT5) of system (1)-(13) in comparison to the solution $u_{0,h}(t) = \int_{\Omega_{\text{cyt}}} u_0(t,x)dx$ of system (14)-(25) (left) and the solution $u_1(t)$ (phosphorylated STAT5) of system (1)-(13) in comparison to the solution $u_{1,h}(t) = \int_{\Omega_{\text{cyt}}} u_1(t,x)dx$ of system (14)-(25) (right) in a NIH3T3-cell.

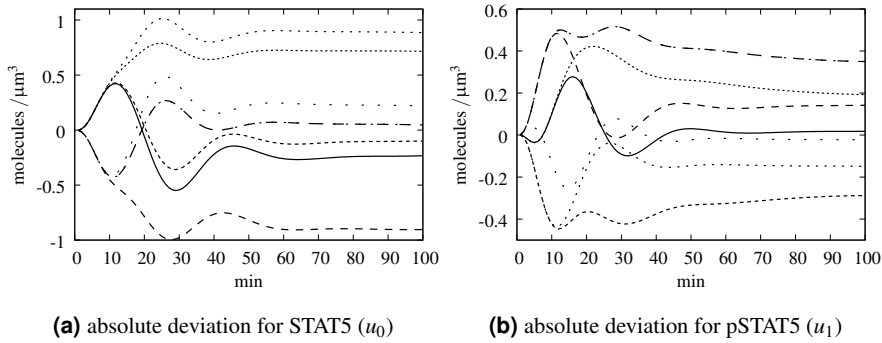


Fig. 7: Sensitivity analysis as a series of tests in which we set different parameter values to see how a change in the parameter causes a change in the dynamic behavior of our model. In this figure we see the absolute deviation between the concentrations of the un- and phosphorylated STAT5 in the cytoplasm of a NIH3T3 cell with rates from table 2.

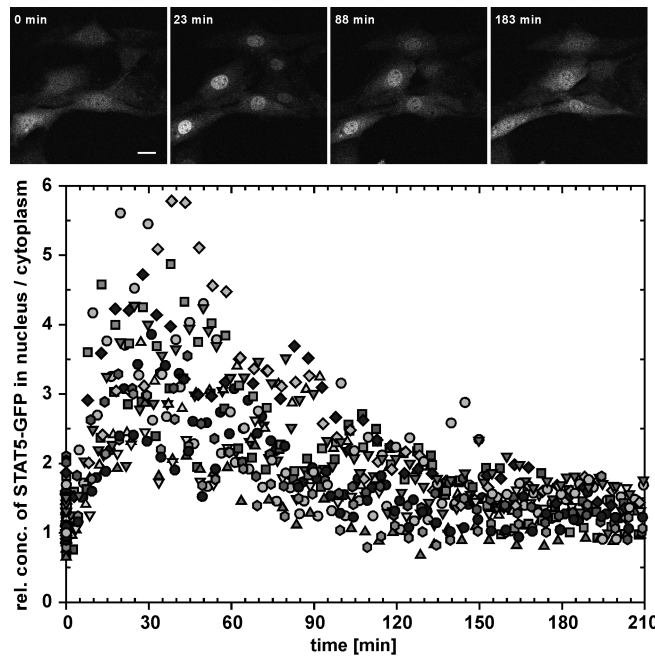


Fig. 8: Live cell imaging of STAT5-GFP nuclear accumulation after Epo stimulation. *NIH3T3-EpoR* cells expressing *STAT5-GFP* were serum-starved for 5 hours and transferred to the microscope. After addition of Epo, cells were imaged on a confocal microscope for a minimum of 4 hours at 37°C. The maximum of nuclear accumulation of *STAT5* was observed around 23 minutes. Scale bar, 20 μm . The graph shows the ratio of total nuclear *STAT5-GFP* to total cytoplasmic *STAT5-GFP* after stimulation with Epo. Different symbols and grey scales represent individual cells.