Systems Biology of the JAK-STAT signalling pathway

Jens Timmer

Center for Systems Biology
Center for Data Analysis and Modeling
Center for Applied Biosciences
Bernstein Center for Computational Neuroscience
Freiburg Institute for Advanced Studies
Department of Mathematics and Physics
University of Freiburg

http://www.fdm.uni-freiburg.de/∼jeti/

Outline

- **Systems Biology**
- **JAK-STAT pathway of the Epo receptor**
- **A dynamical model for JAK-STAT pathway**
- **Observing the unobservable**
- *In silico* **biology: Predicting a new experiment**
- **Infering systems' properties**

Enlarging Physics, Math, Engineering

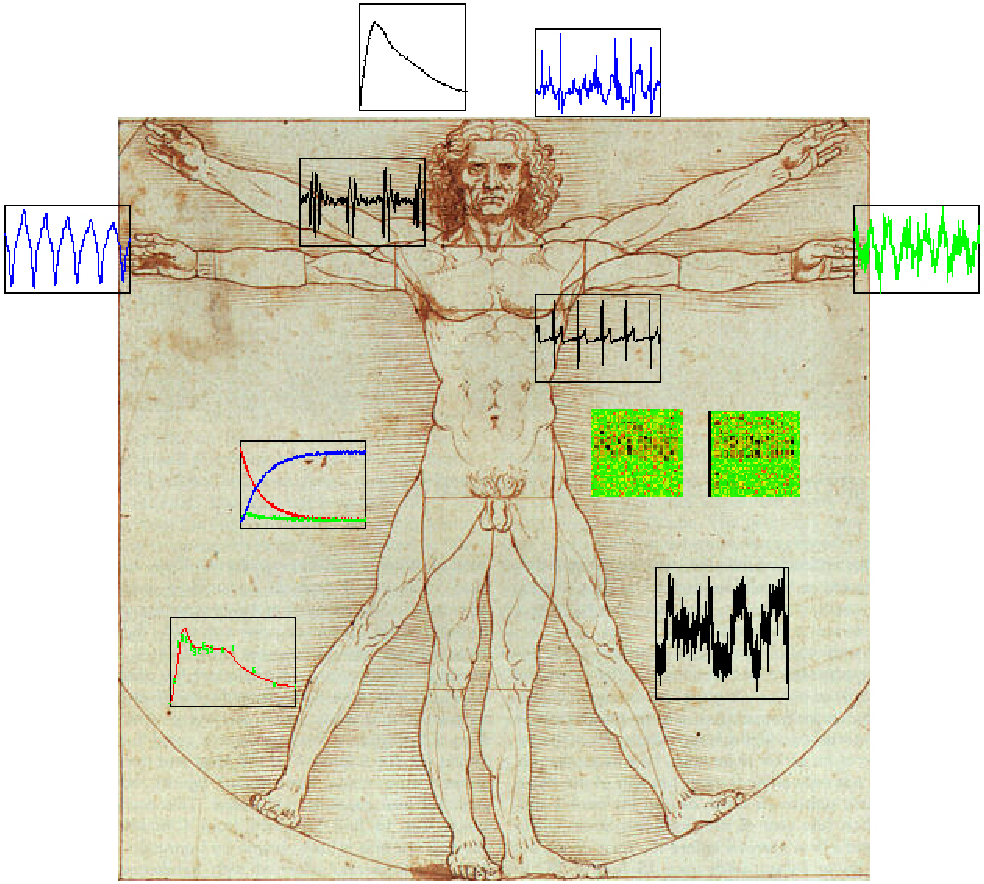
- **Since Newton:**

  Mathematatization of inanimate nature

- **21st century:**

  Additionally: Mathematization of animate nature

Man : A Dynamical System



Diseases caused or expressed by malfunction of dynamical processes

Two Directions in Systems Biology

- **Putting all the omics together**

  So far: large scale, qualitative, static

- **Understanding biomedical networks by data-based mathematical modelling of their dynamical behavior**

  So far: small scale, quantitative, dynamic

Both approaches will converge to: large scale, quantitative, dynamic

Common ground: Investigating networks

Direction II in Systems Biology

Understanding biomedical systems by data-based mathematical modelling of their dynamical behavior

From components and structure to behavior of networks

Systems Biology is based on but more than ...

- **... Mathematical Biology: Data-based**
- **... Bioinformatics: Dynamics**
- **... o.p./g. – o.p.: System**
- **... another omics: Mathematics**

Why Mathematical Modelling in BioMed?

- Make assumptions explicit
- Understand essential properties, failing models
- Condense information, handle complexity
- Understand role of dynamical processes, e.g. feed-back
- Impossible experiments become possible
- Prediction and control
- Understand what is known
- Discover general principles
- "You don't understand it until you can model it"

Why Modelling in Cell Biology?

- Basic Research
  - Genomes are sequenced, but ...
  - ... function determined by regulation
  - Regulation = Interaction & Dynamics
  - Function: Property of dynamic network
  - ”Systems Biology”
- Application
  - Drug development takes 10 years and 1 bn $/€
  - Reduce effort by understanding systems

The (Old) Central Dogma

DNA

$\Downarrow$

RNA

$\Downarrow$

Protein

The (New) Central Dogma

DNA

⇓

RNAs

⇓

Proteins

⇓

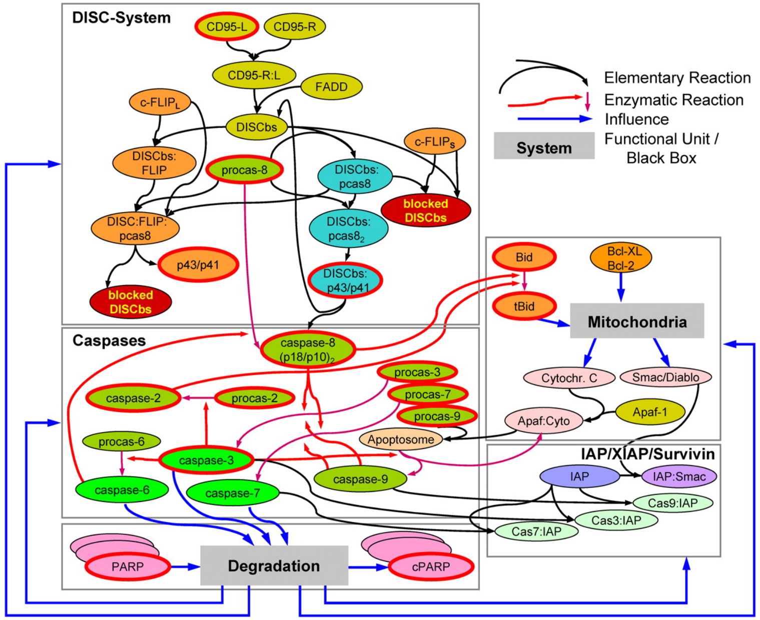
Networks

↙ ↓ ↘

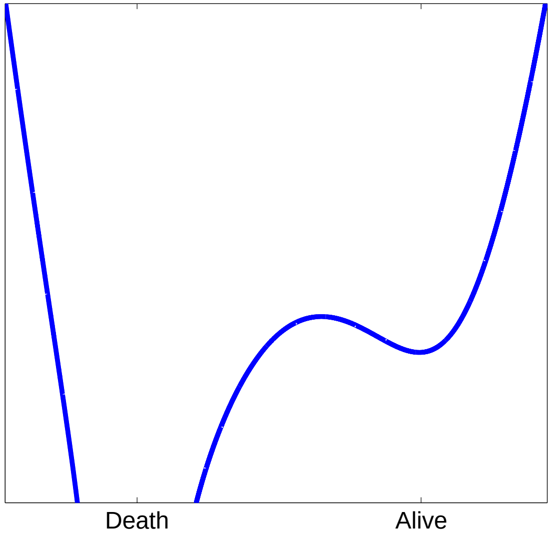
Signalling ↔ **Gene Regulatory** ↔ **Metabolic**

Examples of Networks I: Apoptosis

Pathway cartoon



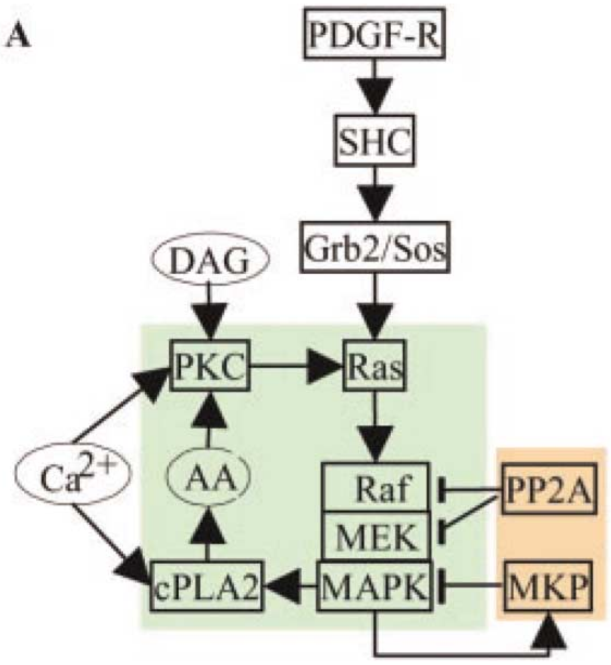
System's behavior



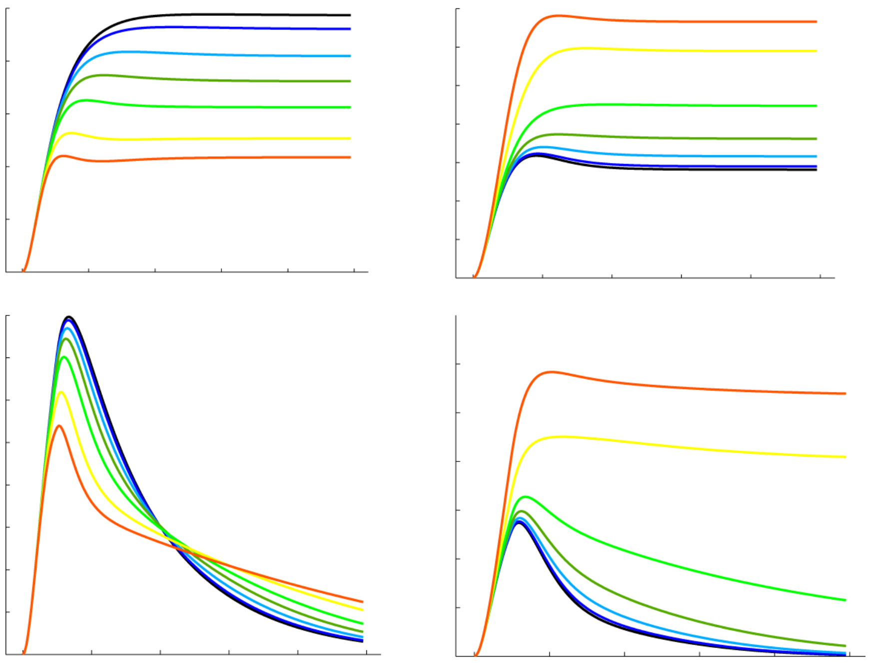
Threshold behavior, one-way bistable

Examples of Networks II: MAP Kinase

Pathway cartoon



System's behavior

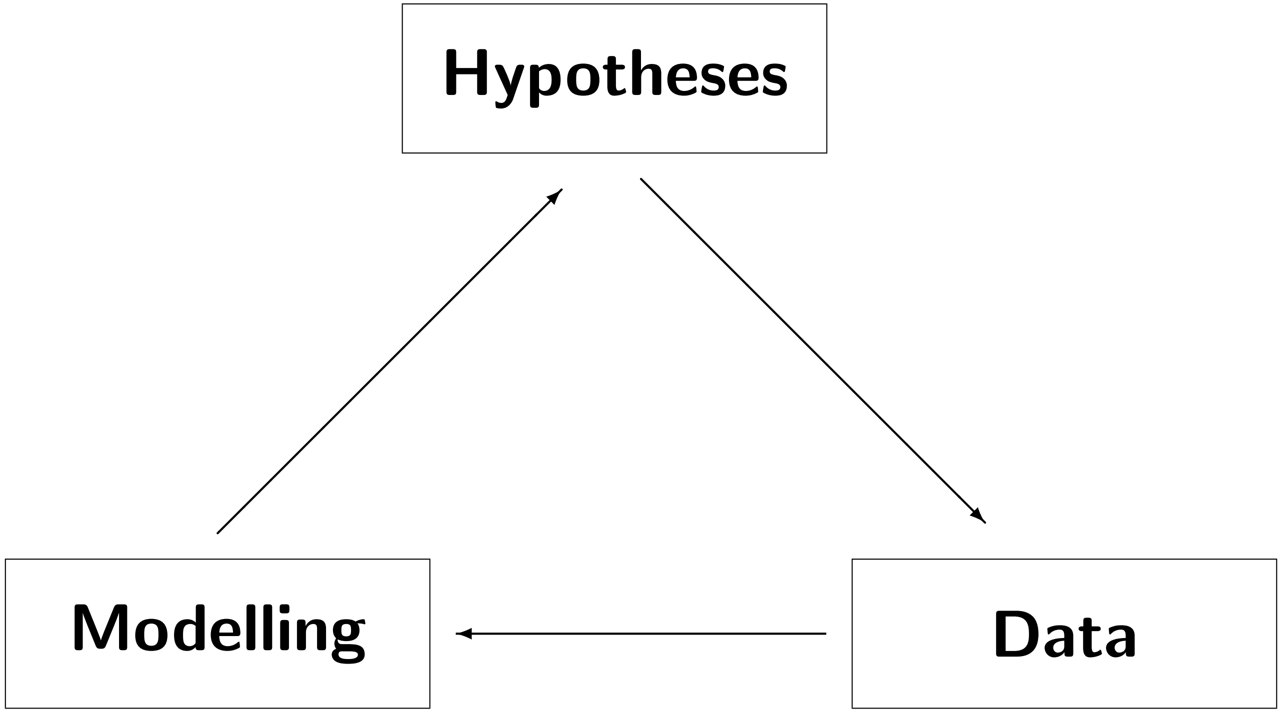


Time scales/parameters important

The Steps of Systems Biology

- **Define the biological question**
- **Modelling**
  - **Experimental design**
  - **Quantitative data**
  - **Parameter estimation**
- **Systems' analysis**
  - **Design priniciples**
  - **Robustness**
- **Applications**
  - **Synthetic biology**
  - **Personalized medicine**

The Systems Biology Cycle: A Process



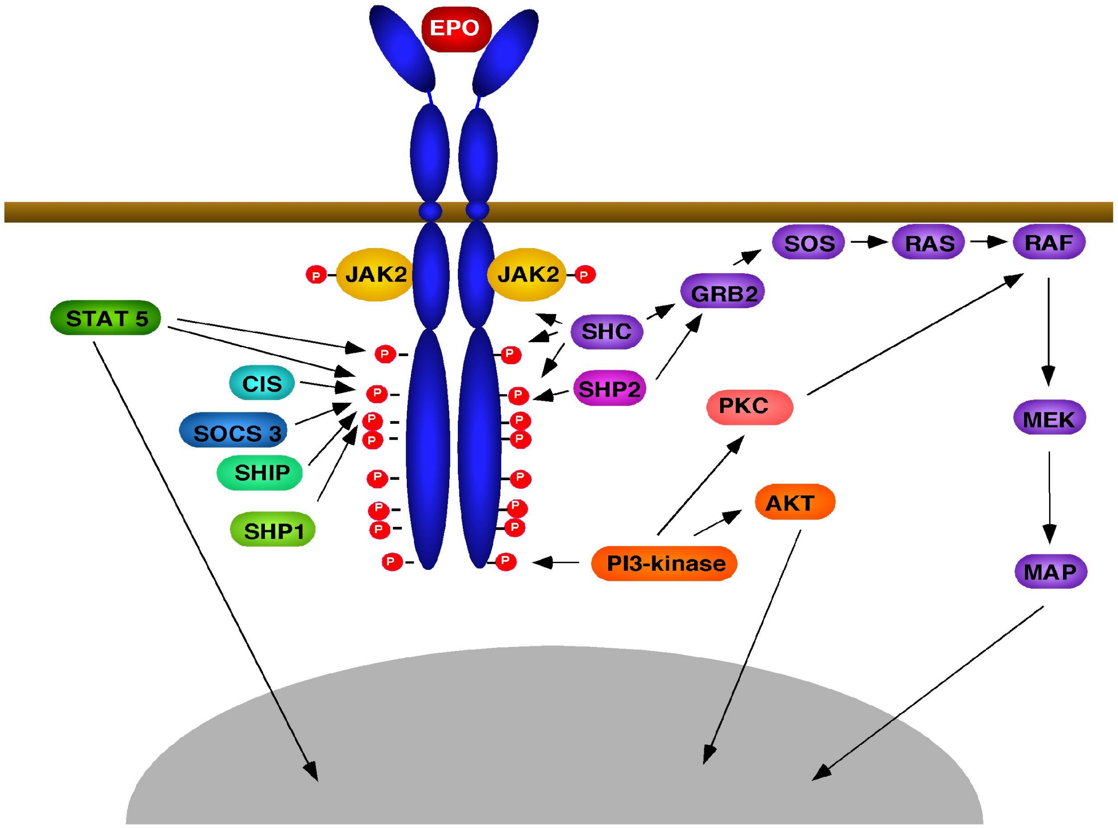
Hypotheses

Modelling

Data

Make the cycle happen: Wet/dry couple projects

Signal transduction through the Erythropoietin receptor (EpoR)



In collaboration with Dr. Ursula Klingmüller

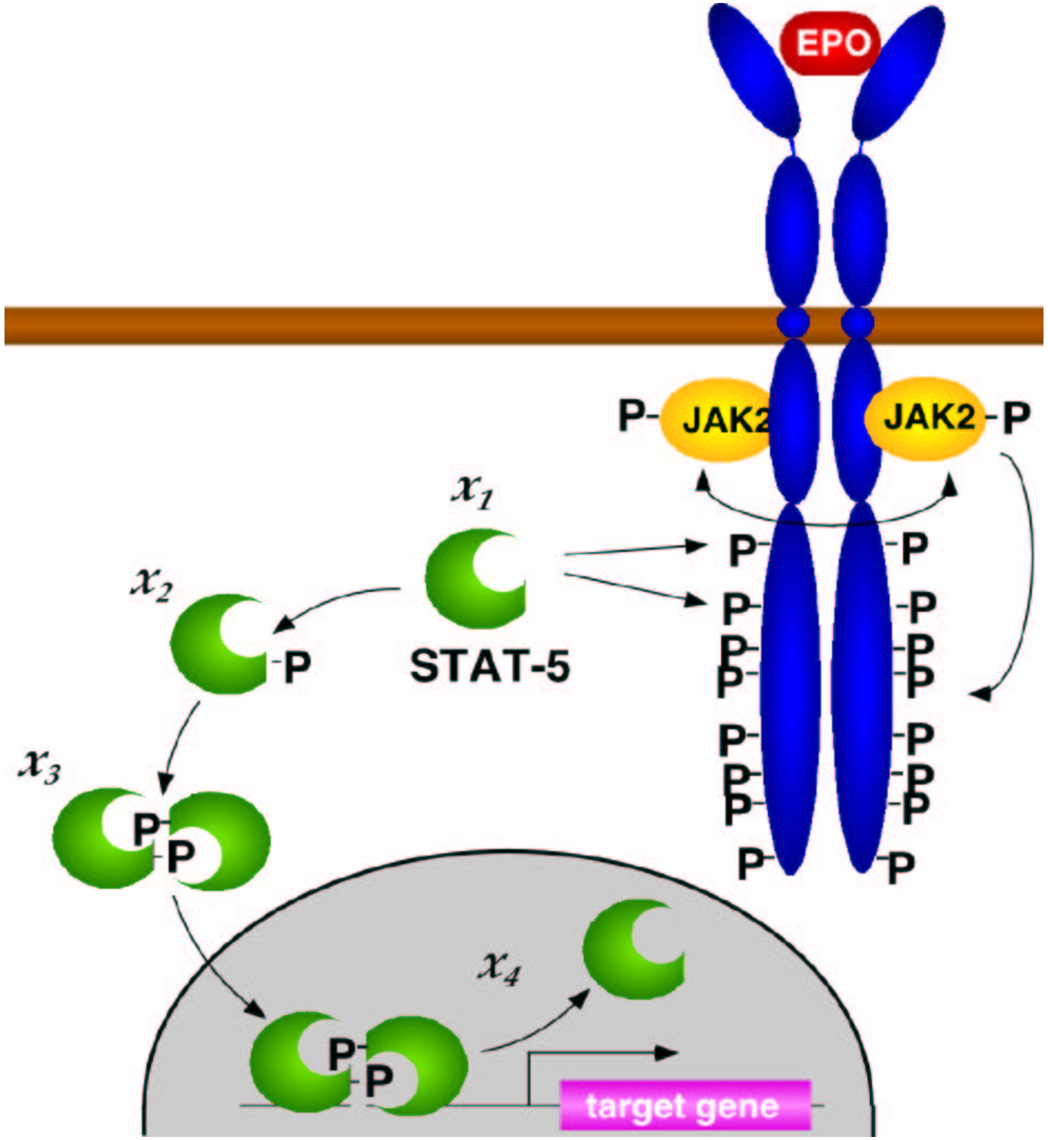
German Cancer Research Centre, Heidelberg

Epo

Epo = Erythropoietin

- **Hormone produced by kidneys**
- **Turns erythroid progenitor cells into red blood cells**
- **Well known to Tour de France cyclists**

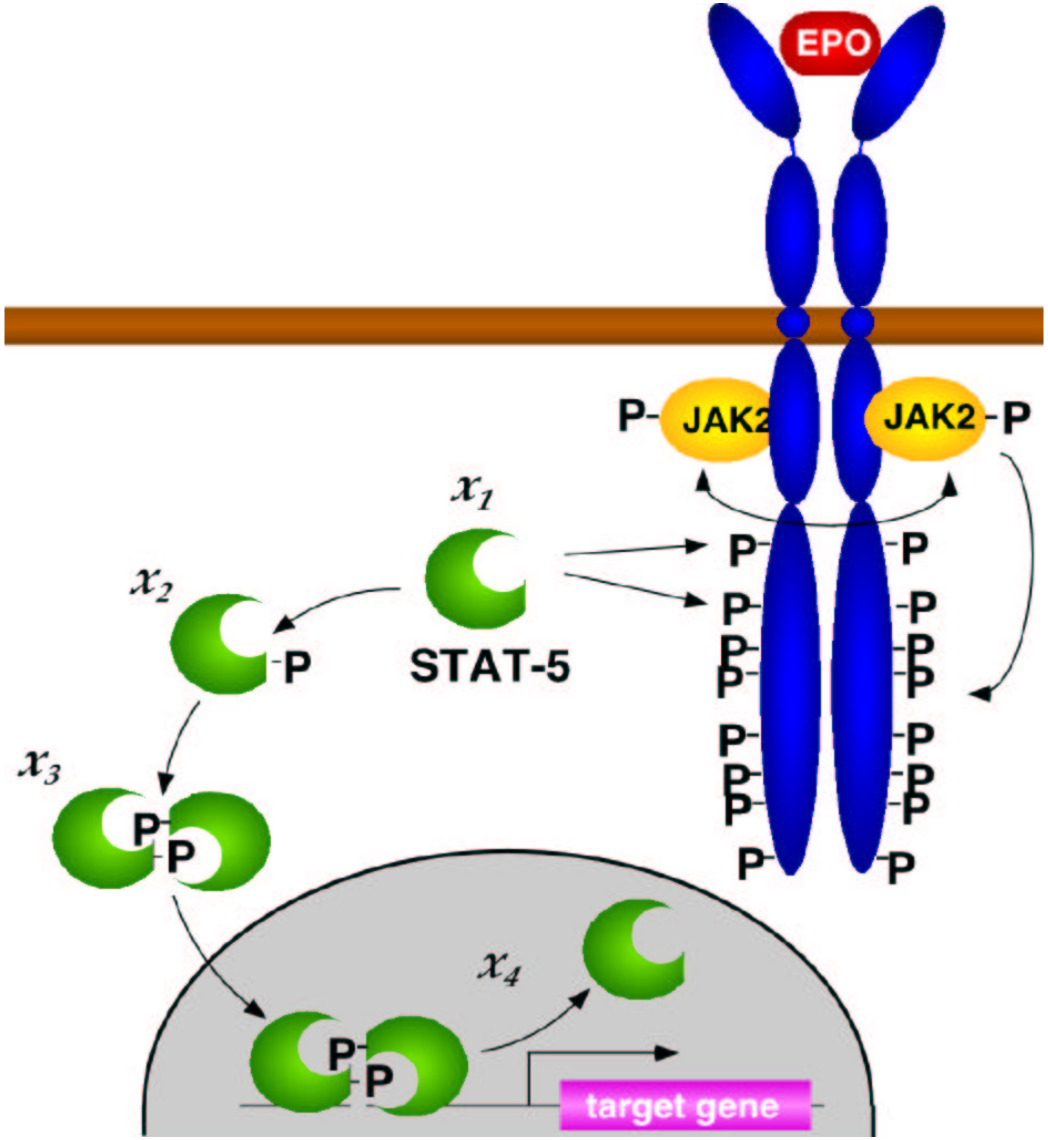
JAK – STAT Pathway



The Program

- Translate the cartoon in (differential) equations
- Measure protein dynamics
- Estimate parameters in equations
- Test and refine the mathematical model
- Predict the outcome of new experiments
- Use the model: E.g. identify potential drug targets

JAK – STAT Pathway



From Chemical Reactions ...

$$\mathbf{STAT5 + EpoR_A \;\rightarrow\; STAT5 - P}$$

$$\mathbf{STAT5 - P + STAT - P \;\rightarrow\; STAT5 - P = STAT5 - P}$$

$$\mathbf{STAT5 - P = STAT5 - P \;\rightarrow\; STAT5 - P = STAT5 - P_{nuc.}}$$

… to Mathematical Equations

$$\begin{aligned}
\dot{x}_1 &= -p_1 x_1 \mathrm{EpoR_A} \\
\dot{x}_2 &= p_1 x_1 \mathrm{EpoR_A} - p_2 x_2^2 \\
\dot{x}_3 &= \frac{1}{2} p_2 x_2^2 - p_3 x_3 \\
\dot{x}_4 &= p_3 x_3
\end{aligned}$$

Measurements

- $\mathbf{y_1(t)}$ **: Phosphorylated STAT-5 in the cytoplasm**

$$\mathbf{y_1(t) = p_5(x_2(t) + 2\,x_3(t))}$$

- $\mathbf{y_2(t)}$ **: All STAT-5 in the cytoplasm**

$$\mathbf{y_2(t) = p_6(x_1(t) + x_2(t) + 2\,x_3(t))}$$

- $\mathbf{y_3(t)}$ **: Activation of the epo receptor**

$$\mathbf{y_3(t) = p_7\,EpoR_A(t)}$$

Simulation vs. Data-Based Modeling I

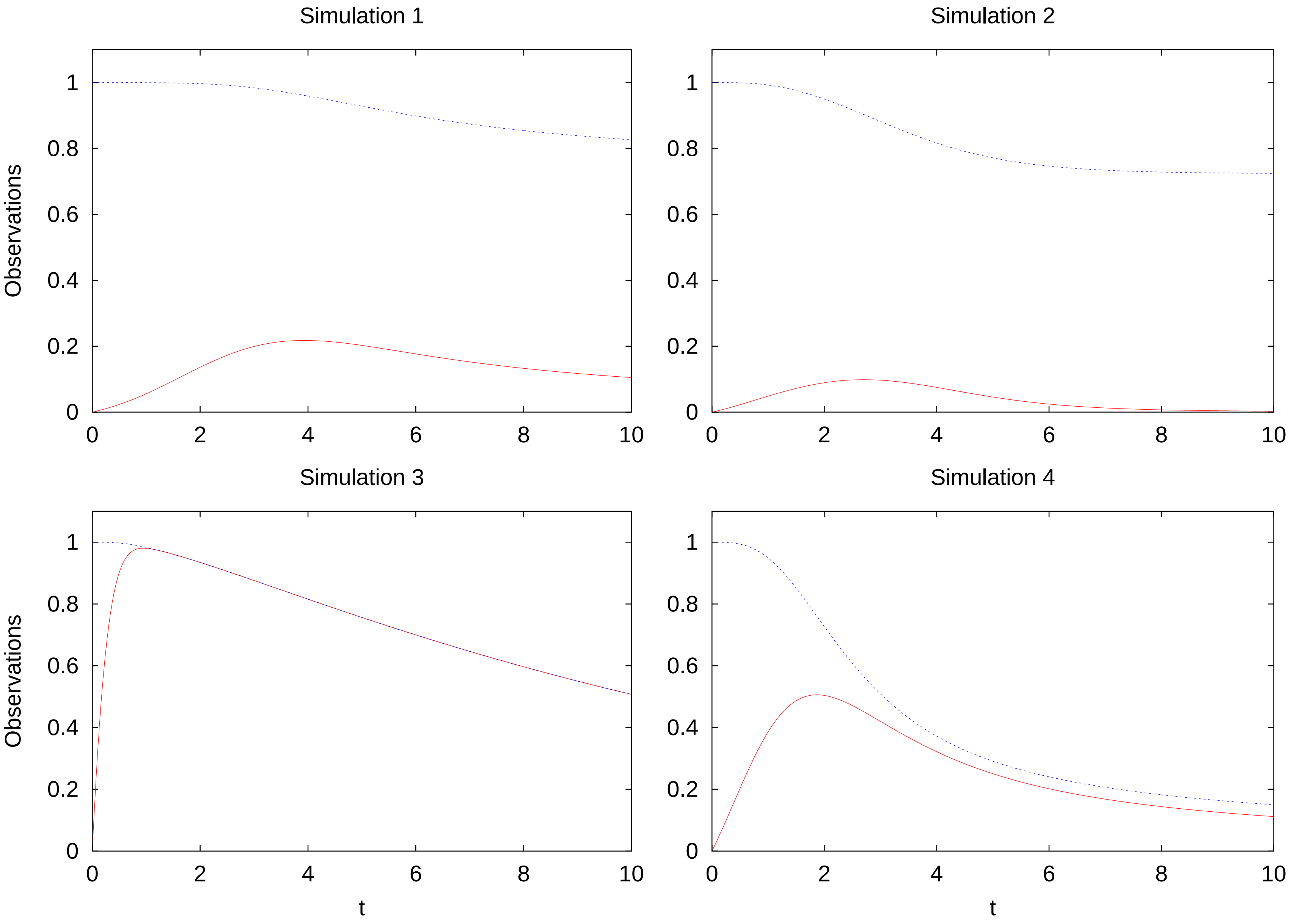
Model comprises:

- **Structure of the equations (the cartoon)**
- **Values of the parameters**

Simulation:

- **Structure from pathway cartoon**
- **Parameters from**
  - **Independent measurements**
  - **Literature**
  - **Educated guesses**

Simulations



Simulation vs. Data-Based Modeling II

Simulation dilemma:

If discrepancies between experiment and model

- **Wrong structure or wrong parameters ?**

Data-based modeling:

- **Structure from pathway cartoon**
- **Parameters estimated from data**

If discrepancies:
Think about the cartoon ! Learn biology !

Parameter Estimation in Nonlinear Partially Observed Noisy Dynamical Systems

Dynamics:

$$\dot{\vec{x}} = \vec{f}(\vec{x}, \vec{p})$$

Observation:

$$\vec{y}(t_i) = \vec{g}(\vec{x}(t_i), \vec{p}) + \vec{\epsilon}(t_i), \qquad \vec{\epsilon}(t_i) \sim N(0, \Sigma_i)$$

Log-Likelihood:

$$E = \chi^2(\vec{p}, \vec{x}(t_0)) = \sum_{i=1}^{N} \sum_{j=1}^{M} \left(\frac{(y_j^D(t_i) - g_j(\vec{x}(t_i; \vec{p}, \vec{x}(t_0))}{\sigma_{i\,j}} \right)^2$$

Really Good Data

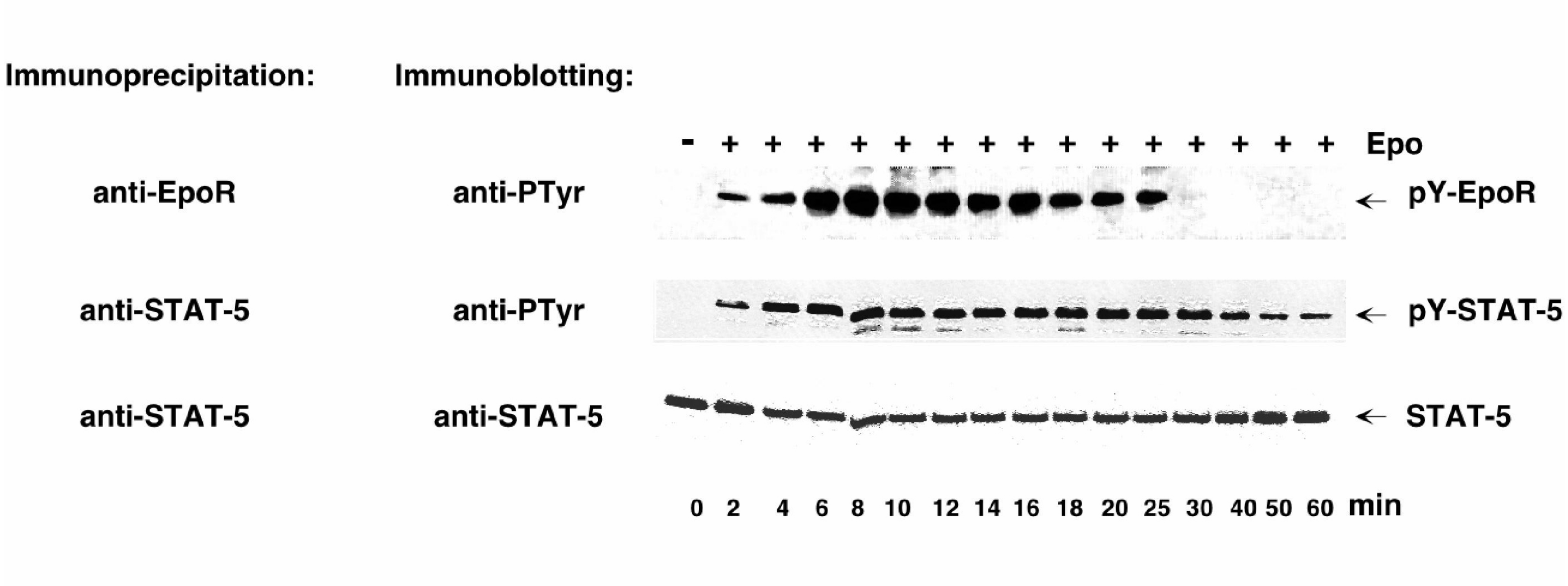
"What makes you feel good ?"

"Good data."

"What makes you feel really good ?"

"Really good data !"

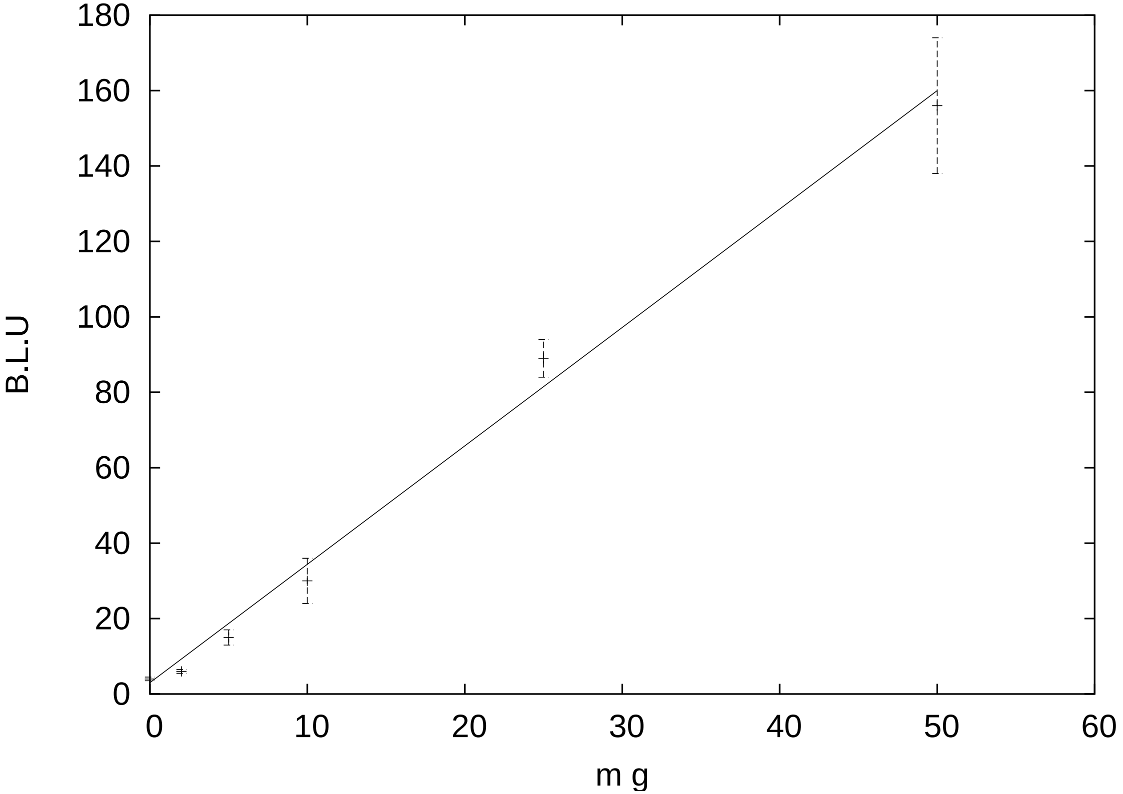
Quantitative Immunoblotting



M. Schilling et al.: Quantitative data generation for Systems Biology. IEE Proc. Sys. Bio. 152, 193, 2005

M. Schilling et al.: Computational processing and error reduction strategies for standardized quantitative data in biological networks. FEBS J. 272, 6400, 2005

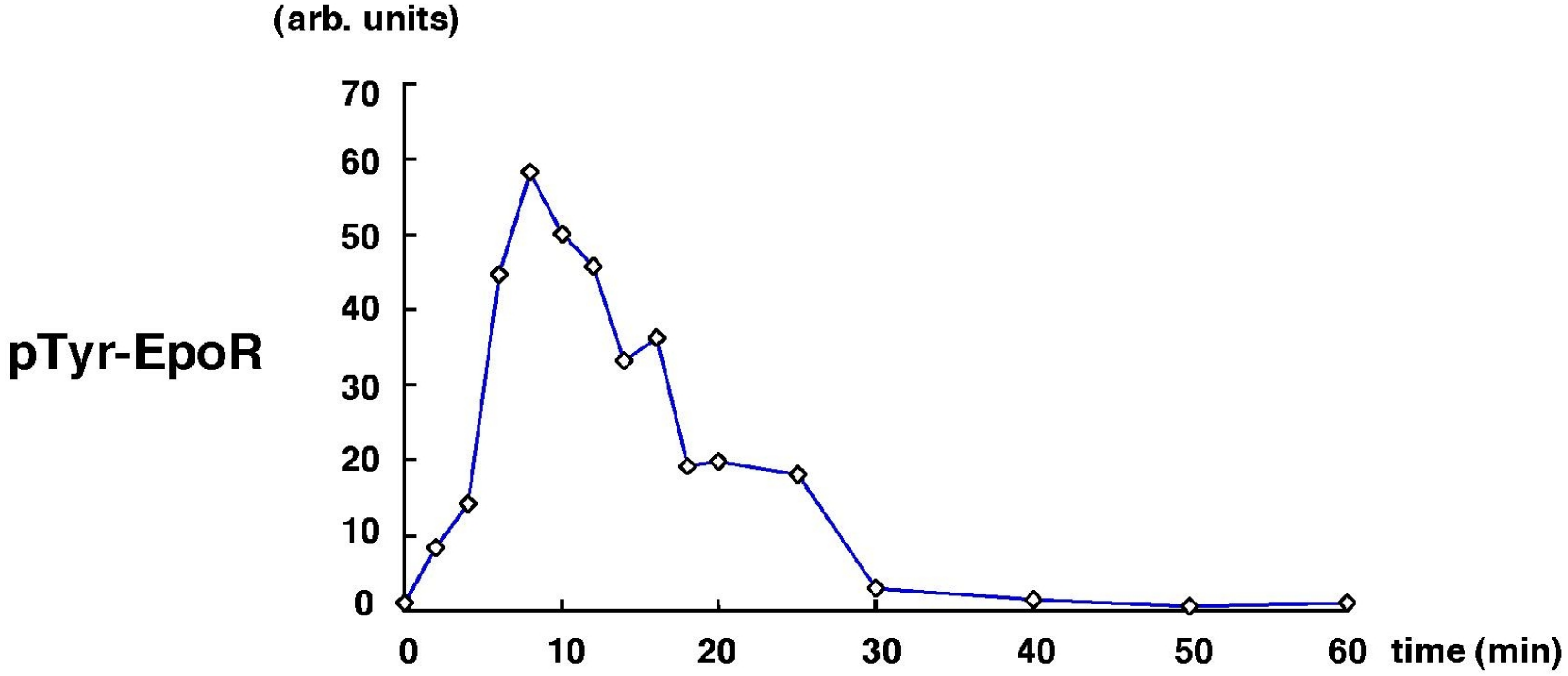
Really Good Data



$g(x)$ is linear

The data

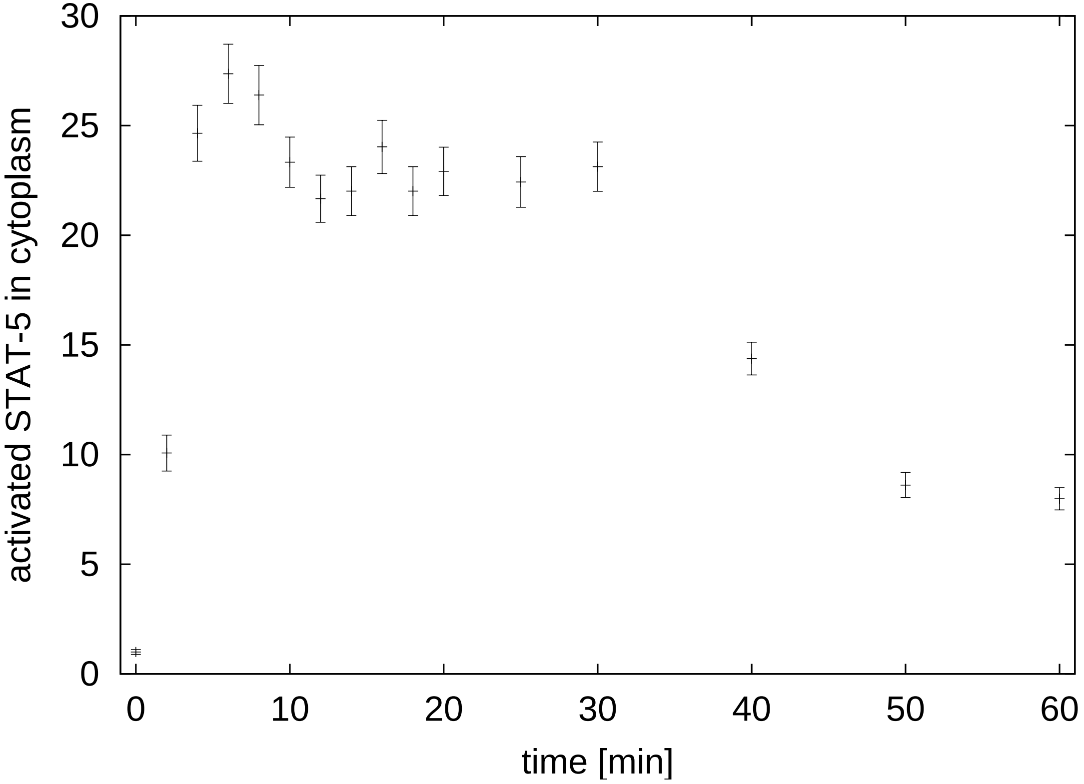
Activation of the epo receptor :



Maximum at 8 min

The data

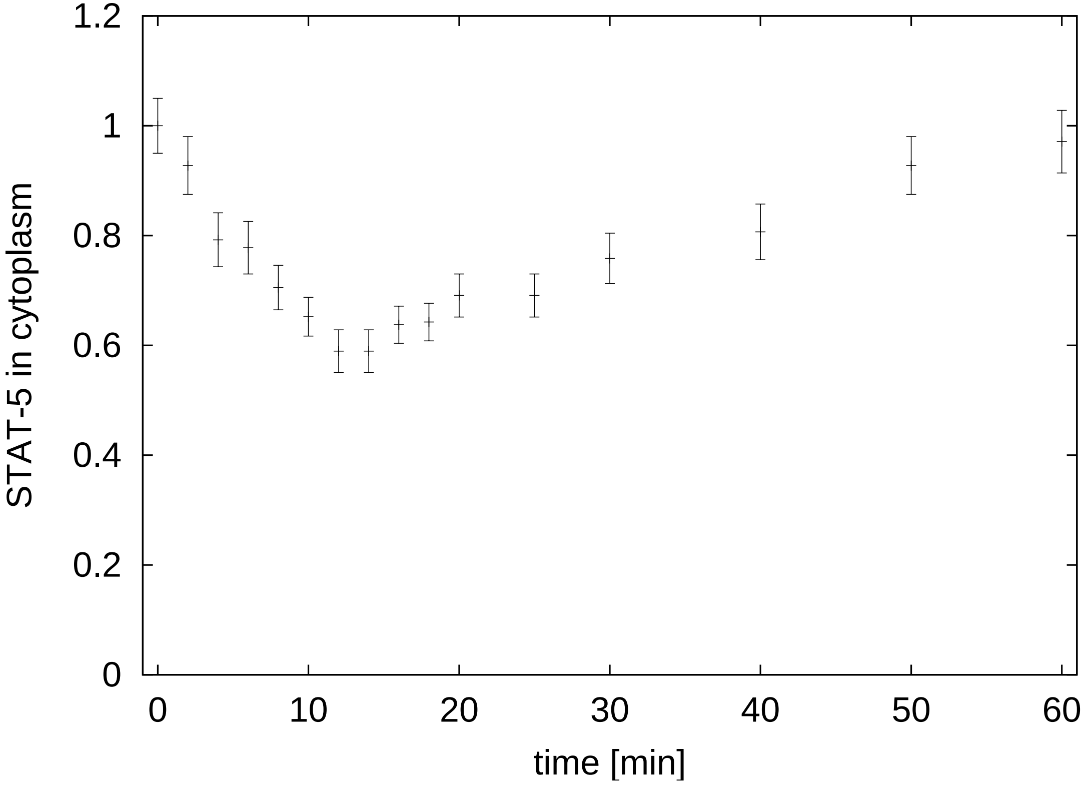
Phosphorylated STAT-5 in cytoplasm :



Plateau from 10 to 30 min

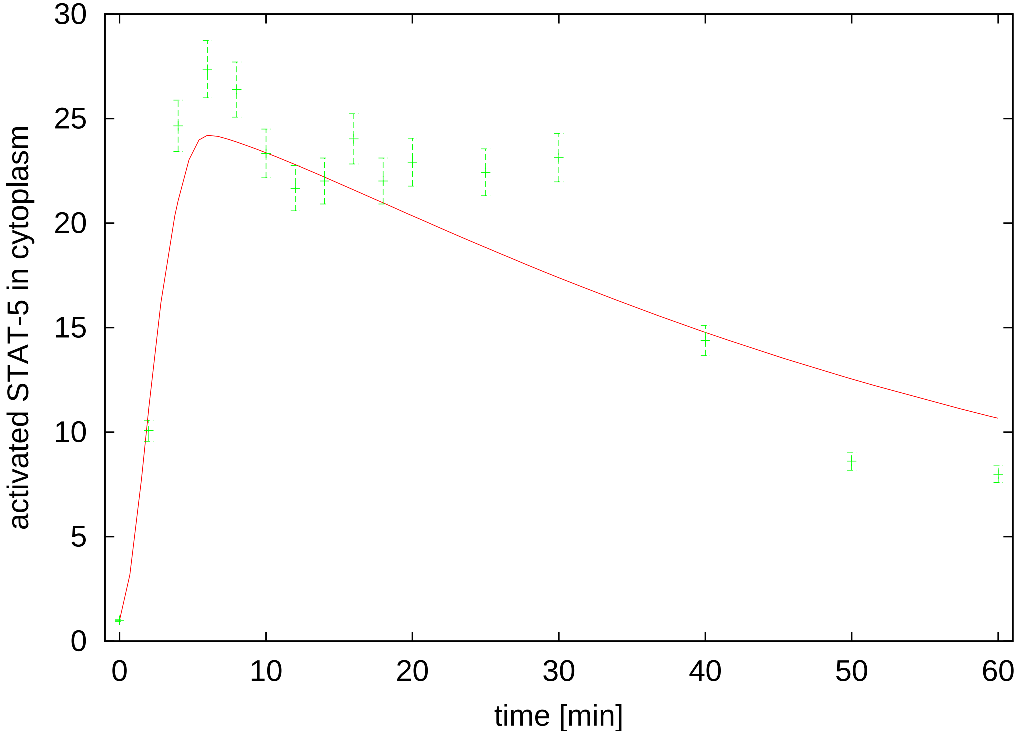
The data

All STAT-5 in cytoplasm :



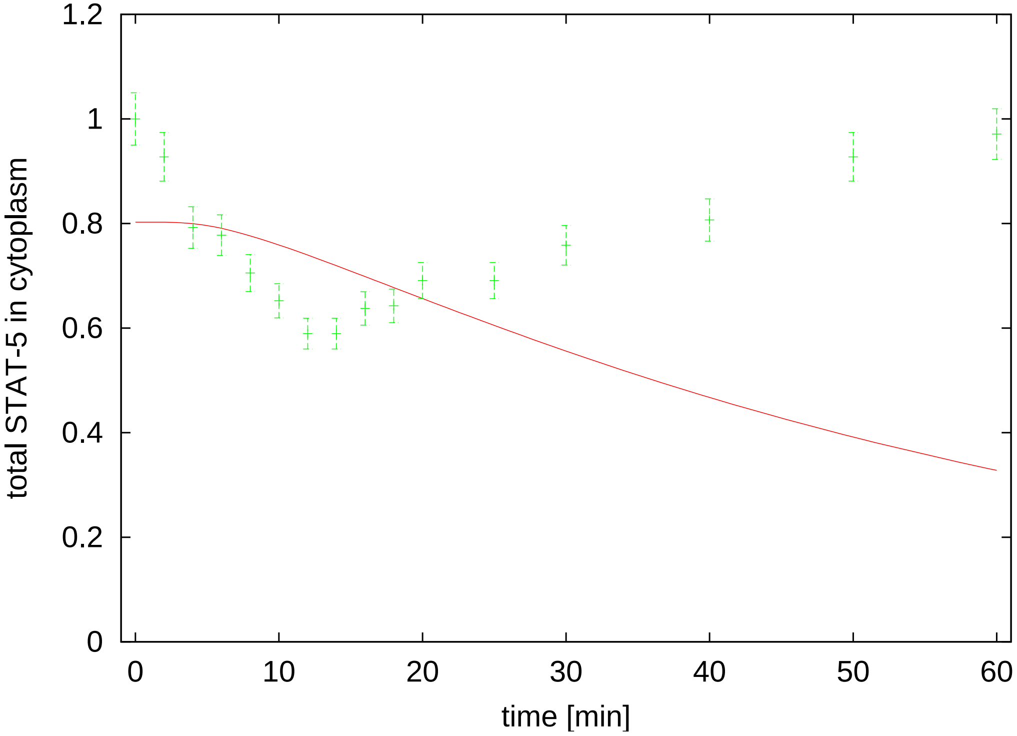
First results

Phosphorylated STAT-5 in cytoplasm :

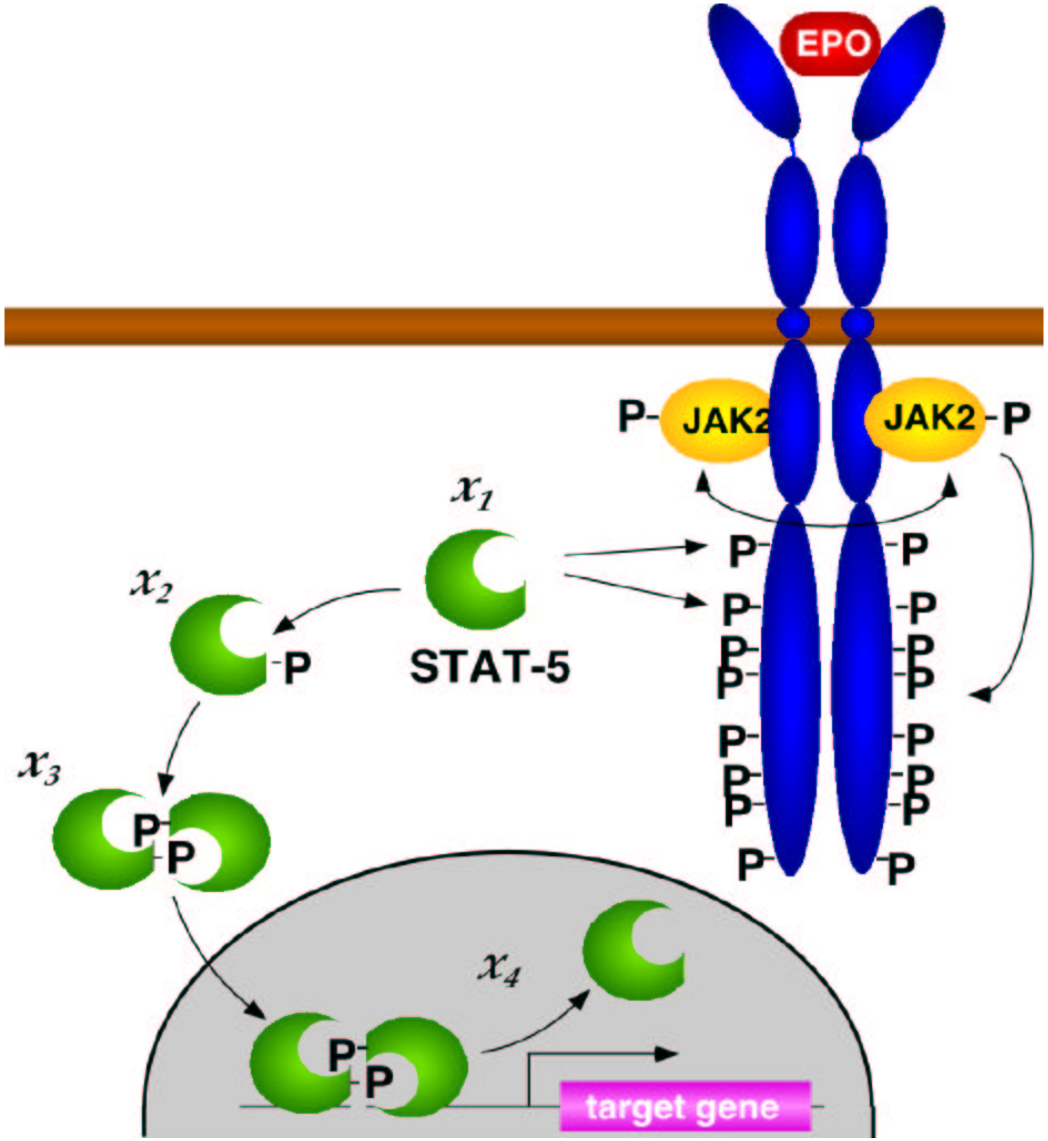


First results

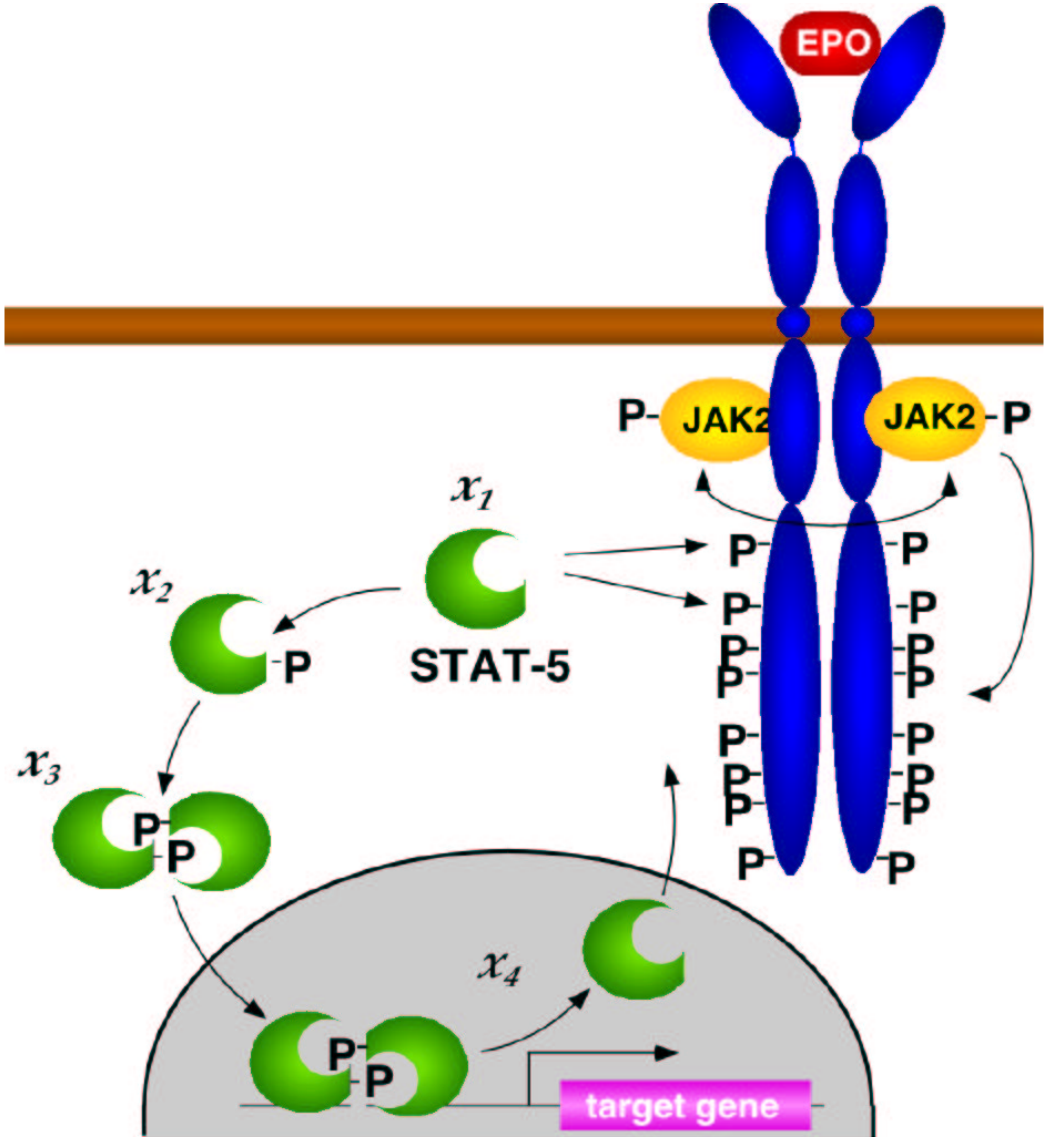
All STAT-5 in cytoplasm :



JAK – STAT Pathway



Model Extension

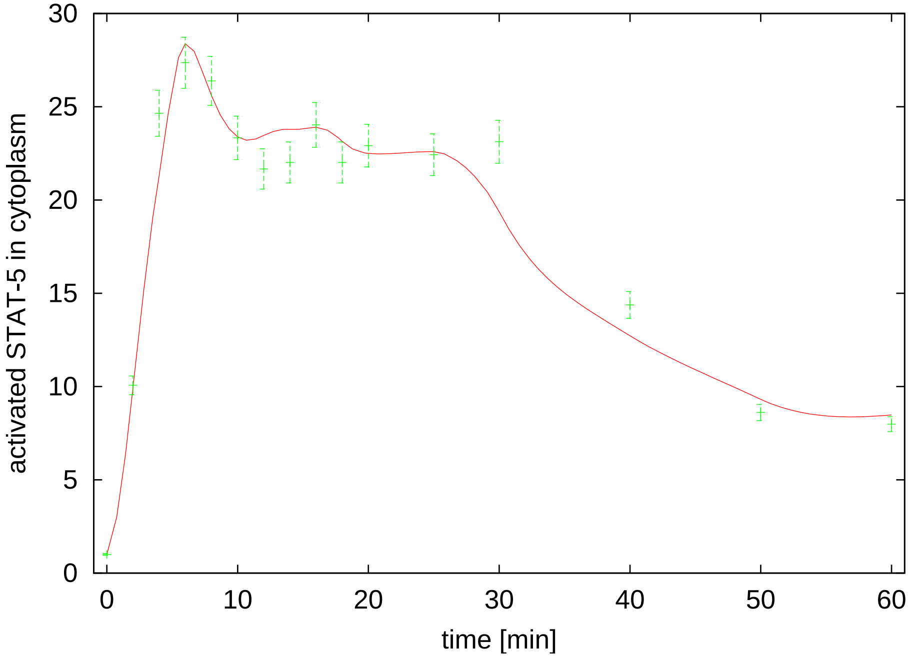


Second try

$$\begin{aligned}
\dot{\mathbf{x}}_1 &= 2\mathbf{p}_4\mathbf{x}_3^{\tau} - \mathbf{p}_1\mathbf{x}_1\mathbf{EpoR}_A \\
\dot{\mathbf{x}}_2 &= \mathbf{p}_1\mathbf{x}_1\mathbf{EpoR}_A - \mathbf{p}_2\mathbf{x}_2^2 \\
\dot{\mathbf{x}}_3 &= \frac{1}{2}\mathbf{p}_2\mathbf{x}_2^2 - \mathbf{p}_3\mathbf{x}_3 \\
\dot{\mathbf{x}}_4 &= \mathbf{p}_3\mathbf{x}_3 - \mathbf{p}_4\mathbf{x}_3^{\tau}
\end{aligned}$$

Results

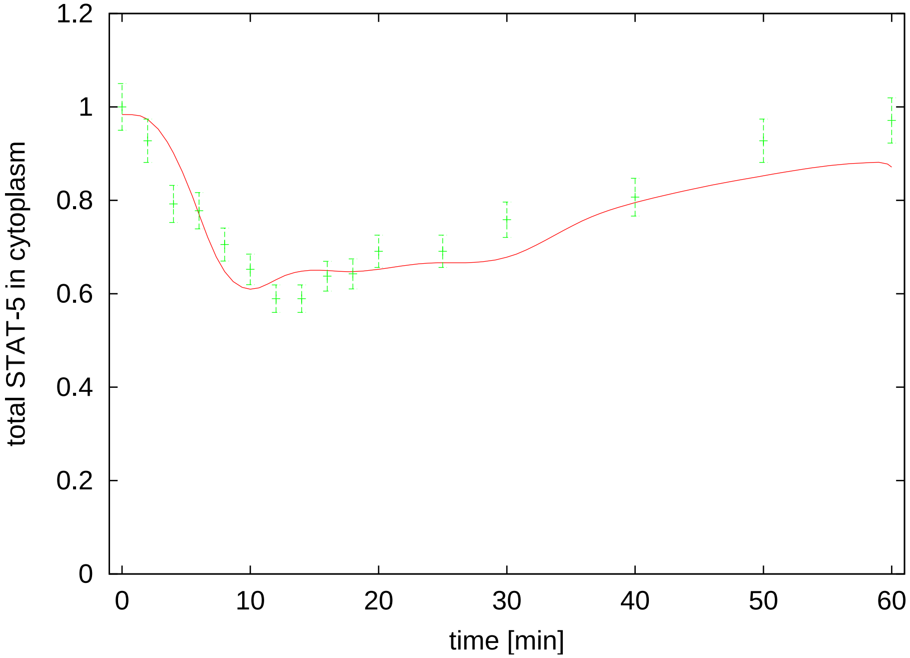
Phosphorylated STAT-5 in cytoplasm :



Sojourn time in nucleus $\tau \approx 6$ min

Results

All STAT-5 in cytoplasm :



Observing the unobservable

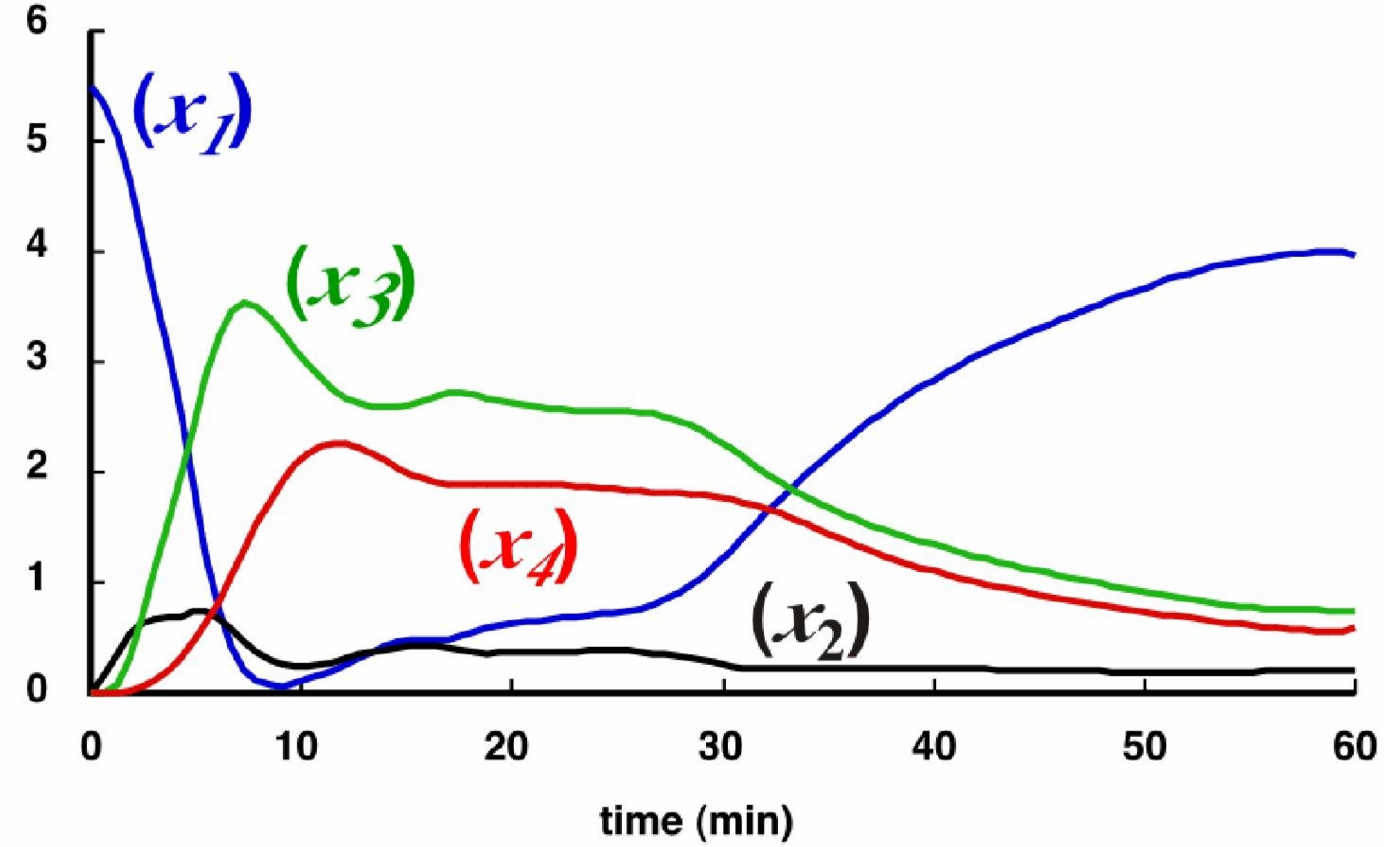
Simulating the fitted model :
Access to dynamic variables x_i

- **Unphophorylated STAT-5 is limiting factor**

- **Experimental fact:**
  Phosphorylated monomeric STAT-5 is hard to measure

  Explanation by the model:
  It is rapidly processed into dimeric STAT-5

Observing the unobservable: The Individual Players


6
5
4
3
2
1
0
(x_1)
(x_3)
(x_4)
(x_2)
0
10
20
30
40
50
60
time (min)

In silico Biology: Impossible Experiments

”What happens if ... ?” Investigations

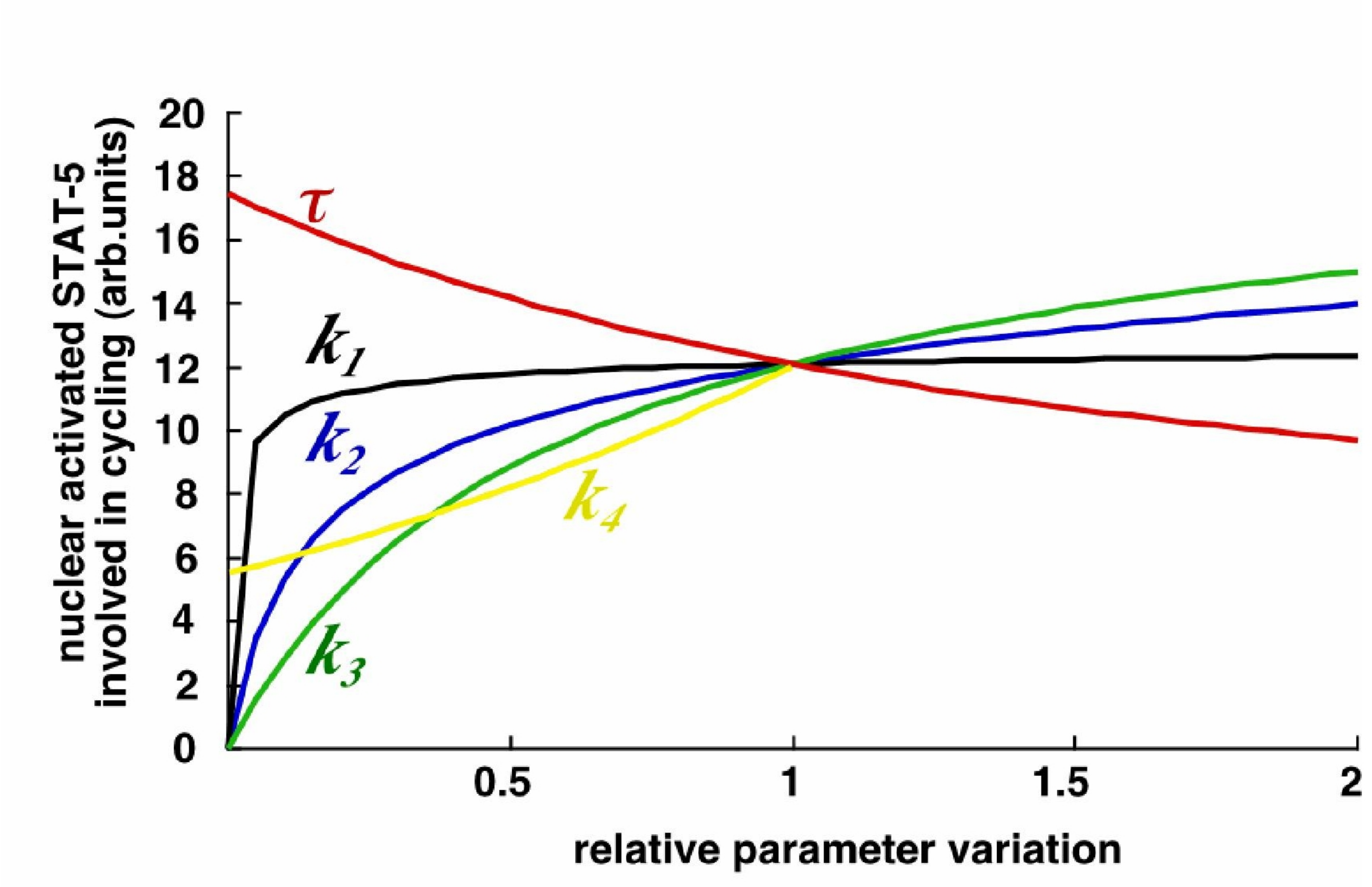
Sensitivity analysis:

- **Change parameters in the model**
- **Calculate the transcriptional yield**

Perspective:

Identification of potential targets for medical intervention

Sensitivity Analysis



Prediction of New Experiment

- Result of sensitivity analysis:

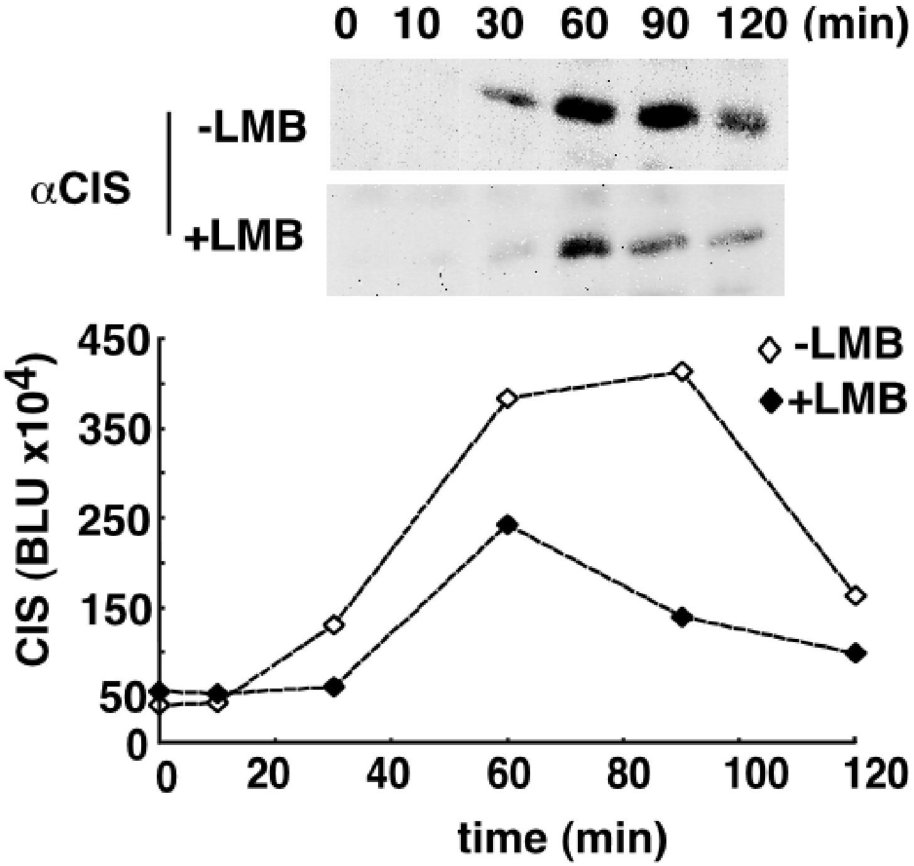
  Transcriptional yield is most sensitive to nuclear shuttling parameters.

- Setting nuclear export to zero

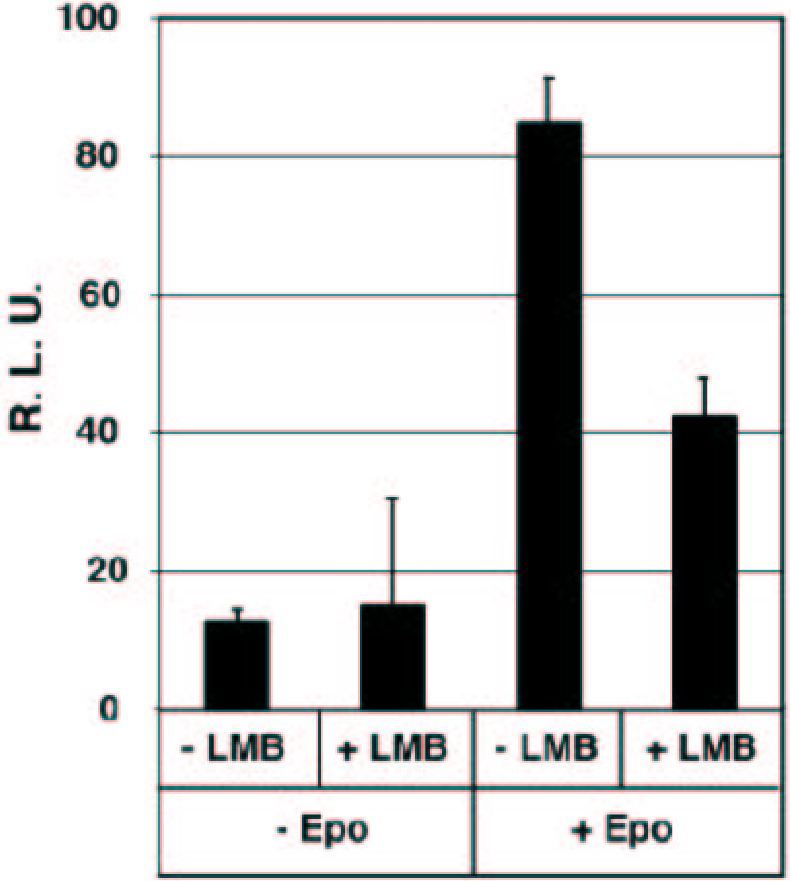
  $\Longrightarrow$ Only one cycle : Only 50 % efficiency

- Blocking nuclear export by leptomycin B confirms prediction.

Experimental Confirmation of Prediction



Experimental Confirmation of Prediction



Why Cycling ?

- **Optimal use of limited pool of STAT-5**
- **Continuous monitoring of receptor activity :**

Systems' property: "Remote Sensor"

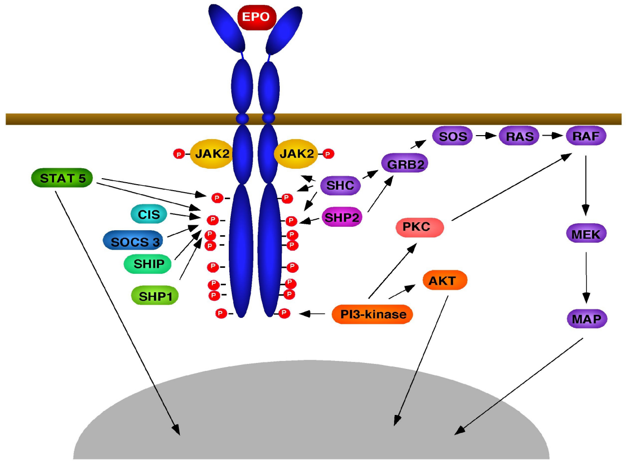
Swameye et al. Proc. Natl. Acad. Sci. 100, 2003, 1028-1033

"All models are wrong ..."

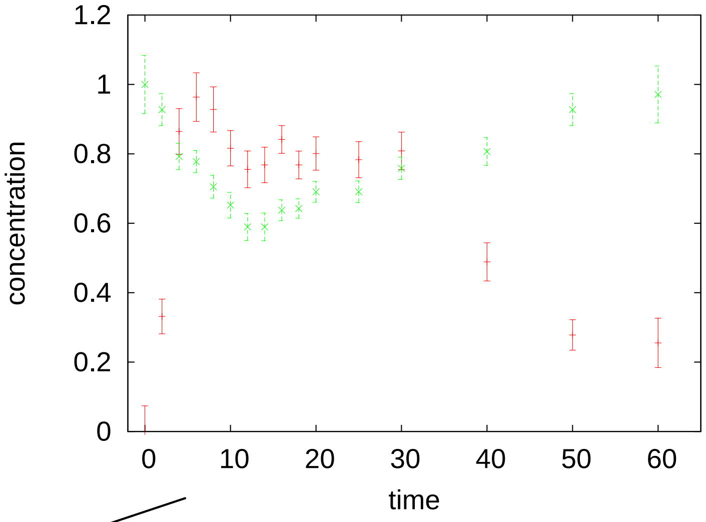
- No scaffolding for receptor–STAT-5 interaction, 200 eqs.
- Spatial effects, ODE vs. PDE
- Stochastic effects
- Data averaged over 10^6 cells

"... but some are useful"

- Captures the main effects
- Makes testable prediction



Signal transduction through the Erythropoietin receptor (EpoR)



$$\dot{\vec{x}} = \vec{f}(\vec{x},\vec{p})$$

In silico **biology**

Test the prior knowledge

Understanding systems' properties

Identification of potential drug targets

Acknowledgements

Theoretical side	Experimental side
	DKFZ, Heidelberg
Thorsten Müller	Ira Swameye
	Olivier Sandra
	Ursula Klingmüller